

Figure 1A

1 CACCCTATCC TACTACTACTA GGAAGCTTGCA CAGTCCGCCT CGGGCAGCCC AAAGCTCCTC 60
 61 TGCCACCCCT GGCTCCCAA ACCCTCCAAA ACAAAGACC AGAAAAGCAC TCTCCACCCA 120
 121 GCAGCCAAAC GCCTCCTTCT TGACGCCAGC CCCCACCCTC TGTCTGCTCG AGCCCAGGAA 180
 181 AGGCCTGAAG GAACAGGCCG GGAAGGAGC CCTCCCTCTC TCCCTTGTC CTCCATCCAC 240
 241 CCAGCGCCGG CATCTGGAGA CCCTATGGCC CGGGCTCACT GGGGCTGCTG CCCCTGGCTG 300
 1 M A R A H W G C C P W L 12
 301 GTCCTCCTCT GTGCTTGTGC CTGGGGCCAC ACAAAGCCAC TGGACCTTGG AGGGCAGGAT 360
 13 V L L C A C A W G H T K P L D L G G Q D 32
 361 GTGAGAAATT GTTCCACCAA CCCCCCTTAC CTTCCAGTTA CTGTGGTCAA TACCACAATG 420
 33 V R N C S T N P P Y L P V T V V N T T M 52
 421 TCACTCACAG CCCTCCGCCA GCAGATGCAG ACCCAGAATC TCTCAGCCTA CATCATCCCA 480
 53 S L T A L R Q Q M Q T Q N L S A Y I I P 72
 481 GGCACAGATG CTCACATGAA CGAGTACATC GGCCAACATG ACGAGAGGCG TGCGTGGATT 540
 73 G T D A H M N E Y I G Q H D E R R A W I 92
 541 ACAGGCTTTA CAGGGTCTGC AGGAACTGCA GTGGTGAATA TGAAGAAAGC AGCTGTCTGG 600
 93 T G F T G S A G T A V V T M K K A A V W 112
 601 ACCGACAGTC GCTACTGGAC TCAGGCTGAG CGGCAAATGG ACTGTAATTG GGAGCTCCAT 660
 113 T D S R Y W T Q A E R Q M D C N W E L H 132
 661 AAGGAAGTTG GCACCACTCC TATGTGCACC TGGCTCCTCA CCGAGATTCC CGCTGGAGGG 720
 133 K E V G T T P I V T W L L T E I P A G G 152
 721 CGTGTGGGTT TTGACCCCTT CCTCTTGTCC ATTGACACCT GGGAGAGTTA TGATCTGGCC 780
 153 R V G F D P F L L S I D T W E S Y D L A 172
 781 CTCCAAGGCT CTAACAGACA GCTGGTGTCC ATCACAACCA ATCTTGTGGA CCTGGTATGG 840
 173 L Q G S N R Q L V S I T T N L V D L V W 192
 841 GGATCAGAGA GGCCACCGGT TCCAAATCAA CCCATTTATG CCCTGCAGGA GGCATTCACA 900
 193 G S E R P P V P N Q P I Y A L Q E A F T 212
 901 GGGAGCACTT GGCAGGAGAA AGTATCTGGC GTCCGAAGCC AGATGCAGAA GCATCAAAAG 960
 213 G S T W Q E K V S G V R S Q M Q K H Q K 232

Figure 1B

```

961 GTCCCGACTG CCGTCCTTCT GTCGGCGCTT GAGGAGACGG CCTGGCTCTT CAACCTTCGA 1020
233 V P T A V L L S A L E E T A W L F N L R 252

1021 GCCAGTGACA TCCCCTATAA CCCCTTCTTC TATTCCTACA CGCTGCTCAC AGACTCTTCT 1080
253 A S D I P Y N P F F Y S Y T L L T D S S 272

1081 ATTAGGTTGT TTGCAAACAA GAGTCGCTTT AGCTCCGAAA CCTTGAGCTA TCTGAAGTCC 1140
273 I R L F A N K S R F S S E T L S Y L N S 292

1141 AGTTGCACAG GCCCCATGTG TGTGCAAATC GAGGATTACA GCCAAGTTCTG TGACAGCATC 1200
293 S C T G P M C V Q I E D Y S Q V R D S I 312

1201 CAGGCCTACT CATGCGGAGA TGTGAGGATC TGGATTGGGA CCAGCTATAC CATGTATGGG 1260
313 Q A Y S L G D V R I W I G T S Y T M Y G 332

1261 ATCTATGAAA TGATACCAAG GGAGAAACTC GTGACAGACA CCTACTCCCC AGTGATGATG 1320
333 I Y E M I P R E K L V T D T Y S P V M M 352

1321 ACCAAGGCAG TGAAGAACAG CAAGGAGCAG GCCCTCCTCA AGGCCAGCCA CGTGCGGGAC 1380
353 T K A V K N S K E Q A L L K A S H V R D 372

1381 GCTGTGGCTG TGATCCGGTA CTTGGTCTGG CTGGAGAAGA ACGTGCCCAA AGGCACAGTG 1440
373 A V A V I R Y L V W L E K N V P K G T V 392

1441 GATGAGTTTT CGGGGGCAGA GATCGTGGAC AAGTTCCGAG GAGAAGAACA GTTCTCCTCC 1500
393 D E F S G A E I V D K F R G E E Q F S S 412

1501 GGACCCAGTT TTGAAACCAT CTCTGCTAGT GGTTTGAATG CTGCCCTGGC CCACTACAGC 1560
413 G P S F E T I S A S G L N A A L A H Y S 432

1561 CCGACCAAGG AGCTGAACCG CAAGCTGTCC TCAGATGAGA TGTACCTGCT GGACTCTGGG 1620
433 P T K E L N R K L S S D E M Y L L D S G 452

1621 GGGCAGTACT GGGACGGGAC CACAGACATC ACCAGAACAG TCCACTGGGG CACCCCCTCT 1680
453 G Q Y W D G T T D I T R T V H W G T P S 472

1681 GCCTTTCAGA AGGAGGCATA TACCCGTGTG CTGATAGGAA ATATTGACCT GTCCAGGCTC 1740
473 A F Q K E A Y T R V L I G N I D L S R L 492

1741 ATCTTTCCCG CTGCTACATC AGGGCGAATG GTGGAGGCCT TTGCCCCGAG AGCCTTGTGG 1800
493 I F P A A T S G R M V E A F A R R A L W 512

1801 GATGCTGGTC TCAATTATGG TCATGGGACA GGCCACGGCA TTGGCAACTT CCTGTGTGTG 1860
513 D A G L N Y G H G T G H G I G N F L C V 532

```

Figure 1C

1861 CATGAGTGGC CAGTGGGATT CCAGTCCAAC AACATCGCTA TGGCCAAGGG CATGTTCACT 1920
 533 H E W P V G F Q S N N I A M A K G M F T 552

1921 TCCATTGAAC CTGGTTACTA TAAGGATGGA GAATTGCGA TCCGTCTCGA AGATGTGGCT 1980
 553 S I E P G Y Y K D G E F G I R L E D V A 572

1981 CTCGTGGTAG AAGCAAAGAC CAAGTACCCA GGGGAGCTAC CTGACCTTGT GGTATCATTT 2040
 573 L V V E A K T K Y P G E L P D L V V S F 592

2041 GTGCCCTATG ACCGGAACCT CATCGATGTC AGCCTGCTGT CTCCCGAGCA TCTCCAGTAC 2100
 593 V P Y D R N L I D V S L L S P E H L Q Y 612

2101 CTGAATCGCT ACTACCAGAC CATCCGGGAG AAGGTGGGTC CAGAGCTGCA GAGGCGCCAG 2160
 613 L N R Y Y Q T I R E K V G P E L Q R R Q 632

2161 CTACTAGAGG AGTTCGAGTG GCTTCAACAG CACACAGAGC CCCTGGCCGC CAGGGCCCCA 2220
 633 L L E E F E W L Q Q H T E P L A A R A P 652

2221 GACACCGCCT CCTGGGCCTC TGTGTTAGTG GTCTCCACCC TTGCCATCCT TGGCTGGAGT 2280
 653 D T A S W A S V L V V S T L A I L G W S 672

2281 GTCTAGAGGC TCCAGACTCT CCTGTTAACC CTCCATCTAG ATGGGGGGCT CCCTTGCTTA 2340
 673 V * 673

2341 GCTCCCCTCA CCCTGCACTG AACATACCCC AAGAGCCCCT GCTGGCCCAT TGCCTAGAAA 2400

2401 CCTTTGCATT CATCCTCCTT CTCCAAGACC TATGGAGAAG GTCCCAGGCC CCAGGAAACA 2460

2461 CAGGGCTTCT TGGCCCCAGA TGGCACCTCC CTGCACCCCG GGGTTGTATA CCACACCCTG 2520

2521 GGCCCCCTAAT CCCAGGCCCC GAAATAGGAA AGCCAGCTAG TCTCTTCTCT TCTGTGATCT 2580

2581 CAGTAGGCCT AACCTATAAC CTAACACAGA CTGCTACAGC TGCTCCCCTC CCGCCAAACA 2640

2641 AAGCCCCAAG AAAACAATGC CCCTACCACC CAAGGGTGCC ATGGTCCCGG GAAAACCCAA 2700

2701 CCTGTCACCG CGTGTGGGCG GTAACCAGAA CTGTTCCCCC CCACCAGGGC TTAAAAATCG 2760

2761 CCCCCACTTT TTAACCATCG TCCATTAACC ACCTGGTGGG CATAGCCAGA GCTGTTTCGAA 2820

2821 CCCAGCCAGG GATGAAAAAT CAACCCCCGA CATGGAACCC ATGATTCCTA AACCCGGGGT 2880

2881 AGGTTCCATG CCAAGTAACA GCAGAGGGAG TTAAGCCATA GGAATTGGC TGTGGAGTAA 2940

Figure 1D

2941 GAGGGAATGC GGTGAGGCAG TGTGGAATAT GACCCTACCA GAGGTTGGAG AACAAACTTG 3000
3001 GGCAGCCGGA ACCCGTCACT ATTTTAGATT CCTGGCATTG GAGGAGCCCT TTGAACTTTC 3060
3061 CAAAGTGCAG CCACAGCTAC AATGCTGTTA AATCCTCCCA CATTTCTTGG ATGCCCCTTC 3120
3121 ACCTTGTTGTG GACAGTGTCT GGTTCCTCCA TTTTACAGAC AGGAAAAGTG AGCTTCAGAC 3180
3181 AGGGGGTGGG CTTTGCCTAA GGACACACAA ATTTGGTTGG GAGTTGATGG GGCCAGATGA 3240
3241 GCCAGCATTG CAGCTGTTTC ACCCTTCAGC AACATGCAGA GTCCCTGAGC CCACCTCCCA 3300
3301 GCCCTCTCCT CATCTCTTGA ACCCACTGTG GTGAGAAGAA TTTGCTCCGG CCAAATTGGC 3360
3361 CGTTAGCCAC CTGGGTCCAC ATCCTGCTAA GACGTTTAAA ACAGCCTAAC AAAGACACTT 3420
3421 GCCTGTGG 3428

Figure 2A

1 CACCCTATCC TACTACTACTA GGAAGCTTGCA CAGTCCGCCT CGGGCAGCCC AAAGCTCCTC 60
 61 TGCCCAACCCT GGCTCCCAA ACCCTCCAAA ACAAAAGACC AGAAAAGCAC TCTCCACCCA 120
 121 GCAGCCAAAC GCCTCCTTCT TGACGCCAGC CCCCACCCTC TGTCTGCTCG AGCCAGGAA 180
 181 AGGCTGAAG GAACAGGCCG GGGAAGGAGC CCTCCCTCTC TCCCTTGTC CTCCATCCAC 240
 241 CCAGCGCCGG CATCTGGAGA CCCTATGGCC CGGGCTCACT GGGGCTGCTG CCCCTGGCTG 300
 1 M A R A H W G C C P W L 12
 301 GTCCTCCTCT GTGCTTGTGC CTGGGGCCAC ACAAGCCAC TGGACCTTGG AGGCAGGAT 360
 13 V L L C A C A W G H T K P L D L G G Q D 32
 361 GTGAGAAATT GTTCCACCAA CCCCCCTTAC CTTCCAGTTA CTGTGGTCAA TACCACAATG 420
 33 V R N C S T N P P Y L P V T V V N T T M 52
 421 TCACTCACAG CCCTCCGCCA GCAGATGCAG ACCCAGAATC TCTCAGCCTA CATCATCCCA 480
 53 S L T A L R Q Q M Q T Q N L S A Y I I P 72
 481 GGCACAGATG CTCACATGAA CGAGTACATC GGCCAACATG ACGAGAGGCG TGCGTGGATT 540
 73 G T D A H M N E Y I G Q H D E R R A W I 92
 541 ACAGGCTTTA CAGGGTCTGC AGGAACTGCA GTGGTGAATA TGAAGAAAGC AGCTGTCTGG 600
 93 T G F T G S A G T A V V T M K K A A V W 112
 601 ACCGACAGTC GCTACTGGAC TCAGGCTGAG CGGCAAATGG ACTGTAATTG GGAGCTCCAT 660
 113 T D S R Y W T Q A E R Q M D C N W E L H 132
 661 AAGGAAGTTG GCACCACTCC TATTGTCACC TGGCTCCTCA CCGAGATTCC CGCTGGAGGG 720
 133 K E V G T T P I V T W L L T E I P A G G 152
 721 CGTGTGGGTT TTGACCCCTT CCTCTTGTC ATTGACACCT GGGAGAGTTA TGATCTGGCC 780
 153 R V G F D P F L L S I D T W E S Y D L A 172
 781 CTCCAAGGCT CTAACAGACA GCTGGTGTCC ATCACAACCA ATCTTGTGGA CCTGGTATGG 840
 173 L Q G S N R Q L V S I T T N L V D L V W 192
 841 GGATCAGAGA GGCCACCGGT TCAAATCAA CCCATTTATG CCCTGCAGGA GGCATTCA 900
 193 G S E R P P V P N Q P I Y A L Q E A F T 212
 901 GGGAGCACTT GGCAGGAGAA AGTATCTGGC GTCCGAAGCC AGATGCAGAA GCATCAAAAG 960
 213 G S T W Q E K V S G V R S Q M Q K H Q K 232

Figure 2B

```

961 GTCCCGACTG CCGTCCTTCT GTCGGCGCTT GAGGAGACGG CCTGGCTCTT CAACCTTCGA 1020
233 V P T A V L L S A L E E T A W L F N L R 252

1021 GCCAGTGACA TCCCCTATAA CCCCTTCTTC TATTCCTACA CGCTGCTCAC AGACTCTTCT 1080
253 A S D I P Y N P F F Y S Y T L L T D S S 272

1081 ATTAGGTTGT TTGCAAACAA GAGTCGCTTT AGCTCCGAAA CCTTGAGCTA TCTGAAGTCC 1140
273 I R L F A N K S R F S S E T L S Y L N S 292

1141 AGTTGCACAG GCCCCATGTG TGTGCAAATC GAGGATTACA GCCAAGTTCG TGACAGCATC 1200
293 S C T G P M C V Q I E D Y S Q V R D S I 312

1201 CAGGCCTACT CATTGGGAGA TGTGAGGATC TGGATTGGGA CCAGCTATAC CATGTATGGG 1260
313 Q A Y S L G D V R I W I G T S Y T M Y G 332

1261 ATCTATGAAA TGATACCAAG GGAGAAATC GTGACAGACA CCTACTCCCC AGTGATGATG 1320
333 I Y E M I P R E K L V T D T Y S P V M M 352

1321 ACCAAGGCAG TGAAGAACAG CAAGGAGCAG GCCCTCCTCA AGGCCAGCCA CGTGCGGGAC 1380
353 T K A V K N S K E Q A L L K A S H V R D 372

1381 GCTGTGGCTG TGATCCGGTA CTTGGTCTGG CTGGAGAAGA ACGTGCCCAA AGGCACAGTG 1440
373 A V A V I R Y L V W L E K N V P K G T V 392

1441 GATGAGTTTT CGGGGGCAGA GATCGTGGAC AAGTTCCGAG GAGAAGAACA GTTCTCCTCC 1500
393 D E F S G A E I V D K F R G E E Q F S S 412

1501 GGACCCAGTT TTGAAACCAT CTCTGCTAGT GGTTTGAATG CTGCCCTGGC CCACTACAGC 1560
413 G P S F E T I S A S G L N A A L A H Y S 432

1561 CCGACCAAGG AGCTGAACCG CAAGCTGTCC TCAGATGAGA TGTACCTGCT GGAAGTCTGG 1620
433 P T K E L N R K L S S D E M Y L L D S G 452

1621 GGGCAGTACT GGGACGGGAC CACAGACATC ACCAGAACAG TCCACTGGGG CACCCCCTCT 1680
453 G Q Y W D G T T D I T R T V H W G T P S 472

1681 GCCTTTCAGA AGGAGGCATA TACCCGTGTG CTGATAGGAA ATATTGACCT GTCCAGGCTC 1740
473 A F Q K E A Y T R V L I G N I D L S R L 492

1741 ATCTTTCCCG CTGCTACATC AGGGCGAATG GTGGAGGCCT TTGCCCCGAG AGCCTTGTGG 1800
493 I F P A A T S G R M V E A F A R R A L W 512

1801 GATGCTGGTC TCAATTATGG TCATGGGACA GGCCACGGCA TTGGCAACTT CCTGTGTGTG 1860
513 D A G L N Y G H G T G H G I G N F L C V 532

```

Figure 2C

1861 CATGAGTGGC CAGTGGGATT CCAGTCCAAC AACATCGCTA TGGCCAAGGG CATGTTCACT 1920
 533 H E W P V G F Q S N N I A M A K G M F T 552

1921 TCCATTGAAC CTGGTTACTA TAAGGATGGA GAATTGGGA TCCGTCTCGA AGATGTGGCT 1980
 553 S I E P G Y Y K D G E F G I R L E D V A 572

1981 CTCGTGGTAG AAGCAAAGAC CAAGTACCCA GGGGAGCTAC CTGACCTTGT GGTATCATTT 2040
 573 L V V E A K T K Y P G E L P D L V V S F 592

2041 GTGCCCTATG ACCGGAACCT CATCGATGTC AGCCTGCTGT CTCCGGAGCA TCTCCAGTAC 2100
 593 V P Y D R N L I D V S L L S P E H L Q Y 612

2101 CTGAATCGCT ACTACCAGAC CATCCGGGAG AAGGTGGGTC CAGAGCTGCA GAGGCGCCAG 2160
 613 L N R Y Y Q T I R E K V G P E L Q R R Q 632

2161 CTA TAGAGG AGTTCGAGTG GCTTCAACAG CACACAGAGC CCCTGGCCGC CAGGGCCCCA 2220
 633 L L E E F E W L Q Q H T E P L A A R A P 652

2221 GACACCGCCT CCTGGGCCTC TGTGTTAGTG GTCTCCACCC TTGCCATCCT TGGCTGGAGT 2280
 653 D T A S W A S V L V V S T L A I L G W S 672

2281 GTCTAGAGGC TCCAGACTCT CCTGTTAACC CTCCATCTAG ATGGGGGGCT CCCTTGCTTA 2340
 673 V * 673

2341 GCTCCCCTCA CCCTGCACTG AACATACCCC AAGAGCCCCT GCTGGCCCAT TGCCTAGAAA 2400

2401 CCTTTGCATT CATCCTCCTT CTCCAAGACC TATGGAGAAG GTCCCAGGCC CCAGGAAACA 2460

2461 CAGGGCTTCT TGGCCCCAGA TGGCACCTCC CTGCACCCCG GGGTTGTATA CCACACCCTG 2520

2521 GGCCCCTAAT CCCAGGCCCC GAAATAGGAA AGCCAGCTAG TCTCTTCTCT TCTGTGATCT 2580

2581 CAGTAGGCCT AACCTATAAC CTAACACAGA CTGCTACAGC TGCTCCCCTC CCGCCAAACA 2640

2641 AAGCCCCAAG AAAACAATGC CCCTACCACC CAAGGGTGCC ATGGTCCCGG GAAAACCCAA 2700

2701 CCTGTCACCG CGTGTGGGC GTAACCAGAA CTGTTCCCCC CCACCAGGGC TTAAAAATCG 2760

2761 CCCCCACTTT TTAACCATCG TCCATTAACC ACCTGGTGGG CATAGCCAGA GCTGTTTCGAA 2820

2821 CCCAGCCAGG GATGAAAAAT CAACCCCCGA CATGGAACCC ATGATTCTTA AACCCGGGGT 2880

2881 AGGTTCCATG CCAAGTAACA GCAGAGGGAG TTAAGCCATA GGAATTGGC TGTGGAGTAA 2940

Figure 2D

2941 GAGGGAATGC GGTGAGGCAG TGTGGAATAT GACCCTACCA GAGGTGGAG AACAACTTG 3000
3001 GGCAGCCGGA ACCCGTCACT ATTTTAGATT CCTGGCATTG GAGGAGCCCT TTGAACTTTC 3060
3061 CAAAGTGCAG CCACAGCTAC AATGCTGTTA AATCCTCCCA CATTCTTGG ATGCCCCCTTC 3120
3121 ACCTTGTTGG GACAGTGTCT GGTTCCTCCA TTTTACAGAC AGGAAAAGTG AGCTTCAGAC 3180
3181 AGGGGGTGGG CTTTGCCTAA GGACACACAA ATTTGGTTGG GAGTTGATGG GGCCAGATGA 3240
3241 GCCAGCATTG CAGCTGTTTC ACCCTTCAGC AACATGCAGA GTCCCTGAGC CCACCTCCCA 3300
3301 GCCCTCTCCT CATTCTCTGA ACCCACTGTG GTGAGAAGAA TTTGCTCCGG CCAAATTGGC 3360
3361 CGTTAGCCAC CTGGGTCCAC ATCCTGCTAA GACGTTTAAA ACAGCCTAAC AAAGACACTT 3420
3421 GCCTGTGG 3428

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 3A

```

1 CTGTGCATGG CATCATCCTG GCCCCTCTA GAGCTCCAAT CCTCCAACCA GAGCCAGCTC 60
1      M A S S W P P L E L Q S S N Q S Q L 18

61 TTCCCTCAAA ATGTACGGC CTGTGACAAT GCTCCAGAAG CCTGGGACCT GCTGCACAGA 120
19 F P Q N A T A C D N A P E A W D L L H R 38

121 GTGCTGCCGA CATTATCAT CTCCATCTGT TTCTTCGGCC TCCTAGGGAA CCTTTTGTGTC 180
39 V L P T F I I S I C F F G L L G N L F V 58

181 CTGTTGGTCT TCCTCCTGCC CCGGCGGCAA CTGAACGTGG CAGAAATCTA CCTGGCCAAC 240
59 L L V F L L P R R Q L N V A E I Y L A N 78

241 CTGGCAGCCT CTGATCTGGT GTTGTCTTTG GGCTTGCCCT TCTGGGCAGA GAATATCTGG 300
79 L A A S D L V F V L G L P F W A E N I W 98

301 AACCAGTTTA ACTGGCCTTT CGGAGCCCTC CTCTGCCGTG TCATCAACGG GGTCACTAAG 360
99 N Q F N W P F G A L L C R V I N G V I K 118

361 GCCAATTTGT TCATCAGCAT CTTCCTGGTG GTGGCCATCA GCCAGGACCG CTACCGCGTG 420
119 A N L F I S I F L V V A I S Q D R Y R V 138

421 CTGGTGACCC CTATGGCCAG CGGAAGGCAG CAGCGGCGGA GGCAGGCCCG GGTACCTGTC 480
139 L V H P M A S G R Q Q R R R Q A R V T C 158

481 GTGCTCATCT GGGTGTGGG GGGCCTCTTG AGCATCCCCA CATTCCTGCT GCGATCCATC 540
159 V L I W V V G G L L S I P T F L L R S I 178

541 CAAGCCGTCC CAGATCTGAA CATCACCGCC TGCATCCTGC TCCTCCCCCA TGAGGCCTGG 600
179 Q A V P D L N I T A C I L L L P H E A W 198

601 CACTTTGCAA GGATTGTGGA GTTAAATATT CTGGGTTTCC TCCTACCACT GGCTGCGATC 660
199 H F A R I V E L N I L G F L L P L A A I 218

661 GTCTTCTTCA ACTACCACAT CCTGGCCTCC CTGCGAACGC GGGAGGAGGT CAGCAGGACA 720
219 V F F N Y H I L A S L R T R E E V S R T 238

721 AGAGTGCGGG GGCCGAAGGA TAGCAAGACC ACAGCGCTGA TCCTCACGCT CGTGGTTGCC 780
239 R V R G P K D S K T T A L I L T L V V A 258

781 TTCCTGGTCT GCTGGGCCCC TTACCACTTC TTTGCCTTCC TGGAATTCTT ATTCCAGGTG 840
259 F L V C W A P Y H F F A F L E F L F Q V 278

841 CAAGCAGTCC GAGGCTGCTT TTGGGAGGAC TTCATTGACC TGGGCCTGCA ATTGGCCAAC 900
279 Q A V R G C F W E D F I D L G L Q L A N 298

```

Figure 3B

```

901 TTCTTTGCCT TCACTAACAG CTCCTGAAT CCAGTAATT ATGTCTTGT GGGCCGGCTC 960
299 F F A F T N S S L N P V I Y V F V G R L 318

961 TTCAGGACCA AGGTCTGGGA ACTTTATAAA CAATGCACCC CTAAAAGTCT TGCTCCAATA 1020
319 F R T K V W E L Y K Q C T P K S L A P I 338

1021 TCTTCATCCC ATAGGAAAGA AATCTTCCAA CTTTCTGGC GGAATTAAAA CAGCATTGAA 1080
339 S S S H R K E I F Q L F W R N * 353

1081 CC 1082
```

1000
900
800
700
600
500
400
300
200
100
0

Figure 4A

```

1 CTGTGCATGG CATCATCTCTG GCCCCTCTA GAGCTCCAAT CCTCCAACCA GAGCCAGCTC 60
1      M A S S W P P L E L Q S S N Q S Q L 18

61 TTCCCTCAAA ATGCTACGGC CTGTGACAAT GCTCCAGAAG CCTGGGACCT GCTGCACAGA 120
19 F P Q N A T A C D N A P E A W D L L H R 38

121 GTGCTGCCGA CATTTATCAT CTCCATCTGT TTCTTCGGCC TCCTAGGGAA CCTTTTGTGTC 180
39 V L P T F I I S I C F F G L L G N L F V 58

181 CTGTTGGTCT TCCTCCTGCC CCGGCGGCAA CTGAACGTGG CAGAAATCTA CCTGGCCAAC 240
59 L L V F L L P R R Q L N V A E I Y L A N 78

241 CTGGCAGCCT CTGATCTGGT GTTTGTCTTG GGCTTGCCCT TCTGGGCAGA GAATATCTGG 300
79 L A A S D L V F V L G L P F W A E N I W 98

301 AACCAGTTTA ACTGGCCTTT CGGAGCCCTC CTCTGCCGTG TCATCAACGG GGTCATCAAG 360
99 N Q F N W P F G A L L C R V I N G V I K 118

361 GCCAATTTGT TCATCAGCAT CTTCTCTGGT GTGGCCATCA GCCAGGACCG CTACCGCGTG 420
119 A N L F I S I F L V V A I S Q D R Y R V 138

421 CTGGTGACCC CTATGGCCAG CGGAAGGCAG CAGCGGCGGA GGCAGGCCCG GGTCACCTGC 480
139 L V H P M A S G R Q Q R R R Q A R V T C 158

481 GTGCTCATCT GGGTTGTGGG GGGCCTCTTG AGCATCCCCA CATTCCTGCT GCGATCCATC 540
159 V L I W V V G G L L S I P T F L L R S I 178

541 CAAGCCGTCC CAGATCTGAA CATCACC GCC TGCATCCTGC TCCTCCCCCA TGAGGCCTGG 600
179 Q A V P D L N I T A C I L L L P H E A W 198

601 CACTTTGCAA GGATTGTGGA GTTAAATATT CTGGGTTTCC TCCTACCACT GGCTGCGATC 660
199 H F A R I V E L N I L G F L L P L A A I 218

661 GTCTTCTTCA ACTACCACAT CCTGGCCTCC CTGCGAACGC GGGAGGAGGT CAGCAGGACA 720
219 V F F N Y H I L A S L R T R E E V S R T 238

721 AGAGTGC GGGG GCGCAAGGA TAGCAAGACC ACAGCGCTGA TCCTCACGCT CGTGGTTGCC 780
239 R V R G P K D S K T T A L I L T L V V A 258

781 TTCCTGGTCT GCTGGGCCCC TTACCACTTC TTTGCCTTCC TGGAATTCTT ATTCCAGGTG 840
259 F L V C W A P Y H F F A F L E F L F Q V 278

841 CAAGCAGTCC GAGGCTGCTT TTGGGAGGAC TTCATGACC TGGGCCTGCA ATTGGCCAAC 900
279 Q A V R G C F W E D F I D L G L Q L A N 298

```

Figure 4B

901 TTCTTTGCCT TCACTAACAG CTCCTGAAT CCAGTAATT ATGTCTTTGT GGGCCAGCTC 960
299 F F A F T N S S L N P V I Y V F V G Q L 318

961 TTCAGGACCA AGGTCTGGGA ACTTTATAAA CAATGCACCC CTAAAAGTCT TGCTCCAATA 1020
319 F R T K V W E L Y K Q C T P K S L A P I 338

1021 TCTTCATCCC ATAGGAAAGA AATCTTCCAA CTTTCTGGC GGAATTAAAA CAGCATTGAA 1080
339 S S S H R K E I F Q L F W R N * 353

1081 CC 1082

1081 CC 1082

Figure 5A

```

1 CTGTGCATGG CATCATCCTG GCCCCCTCTA GAGCTCCAAT CCTCCAACCA GAGCCAGCTC 60
1      M A S S W P P L E L Q S S N Q S Q L 18

61 TTCCCTCAAA ATGCTACGGC CTGTGACAAT GCTCCAGAAG CCTGGGACCT GCTGCACAGA 120
19 F P Q N A T A C D N A P E A W D L L H R 38

121 GTGCTGCCAA CATTTATCAT CTCCATCTGT TTCTTCGGCC TCCTAGGGAA CCTTTTGTGTC 180
39 V L P T F I I S I C F F G L L G N L F V 58

181 CTGTTGGTCT TCCTCCTGCC CCGGCGGCAA CTGAACGTGG CAGAAATCTA CCTGGCCAAC 240
59 L L V F L L P R R Q L N V A E I Y L A N 78

241 CTGGCAGCCT CTGATCTGGT GTTTGTCTTG GGCTTGCCCT TCTGGGCAGA GAATATCTGG 300
79 L A A S D L V F V L G L P F W A E N I W 98

301 AACCAGTTTA ACTGGCCTTT CGGAGCCCTC CTCTGCCGTG TCATCAACGG GGTCATCAAG 360
99 N Q F N W P F G A L L C R V I N G V I K 118

361 GCCAATTTGT TCATCAGCAT CTTCTCTGGT GTGGCCATCA GCCAGGACCG CTACCGCGTG 420
119 A N L F I S I F L V V A I S Q D R Y R V 138

421 CTGGTGCACC CTATGGCCAG CGGAAGGCAG CAGCGGCGGA GGCAGGCCCG GGTCACCTGC 480
139 L V H P M A S G R Q Q R R R Q A R V T C 158

481 GTGCTCATCT GGGTTGTGGG GGGCCTCTTG AGCATCCCCA CATTCCTGCT GCGATCCATC 540
159 V L I W V V G G L L S I P T F L L R S I 178

541 CAAGCCGTCC CAGATCTGAA CATCACC GCC TGCATCCTGC TCCTCCCCCA TGAGGCCTGG 600
179 Q A V P D L N I T A C I L L L P H E A W 198

601 CACTTTGCAA GGATTGTGGA GTTAAATATT CTGGGTTTCC TCCTACCACT GGCTGCGATC 660
199 H F A R I V E L N I L G F L L P L A A I 218

661 GTCTTCTTCA ACTACCACAT CCTGGCCTCC CTGCGAACGC GGGAGGAGGT CAGCAGGACA 720
219 V F F N Y H I L A S L R T R E E V S R T 238

721 AGAGTGCGGG GGCCGAAGGA TAGCAAGACC ACAGCGCTGA TCCTCACGCT CGTGGTTGCC 780
239 R V R G P K D S K T T A L I L T L V V A 258

781 TTCCTGGTCT GCTGGGCCCC TTACCACTTC TTTGCCTTCC TGGAAATCTT ATTCCAGGTG 840
259 F L V C W A P Y H F F A F L E F L F Q V 278

841 CAAGCAGTCC GAGGCTGCTT TTGGGAGGAC TTCATTGACC TGGGCCTGCA ATTGGCCAAC 900
279 Q A V R G C F W E D F I D L G L Q L A N 298

```

Figure 5B

```

901 TTCTTTGCCT TCACTAACAG CTCCTGAAT CCAGTAATT ATGTCTTGT GGGCCGGCTC 960
299 F F A F T N S S L N P V I Y V F V G R L 318

961 TTCAGGACCA AGGTCTGGGA ACTTTATAAA CAATGCACCC CTAAAAGTCT TGCTCCAATA 1020
319 F R T K V W E L Y K Q C T P K S L A P I 338

1021 TCTTCATCCC ATAGGAAAGA AATCTTCCAA CTTTCTGGC GGAATTAAAA CAGCATTGAA 1080
339 S S S H R K E I F Q L F W R N * 353

1081 CC 1082
```

1081 CC 1082

Figure 6A

```

1 ATGTTCTCTC CCTGGAAGAT ATCAATGTTT CTGTCTGTTC GTGAGGACTC CGTGCCCAACC 60
1 M F S P W K I S M F L S V R E D S V P T 20

61 ACGGCCTCTT TCAGCGCCGA CATGCTCAAT GTCACCTTGC AAGGGCCAC TCTTAACGGG 120
21 T A S F S A D M L N V T L Q G P T L N G 40

121 ACCTTTGCCC AGAGCAAATG CCCCCAAGTG GAGTGGCTGG GCTGGCTCAA CACCATCCAG 180
41 T F A Q S K C P Q V E W L G W L N T I Q 60

181 CCCCCCTTCC TCTGGGTGCT GTTCGTGCTG GCCACCCTAG AGAACATCTT TGTCTCAGC 240
61 P P F L W V L F V L A T L E N I F V L S 80

241 GTCTTCTGCC TGCACAAGAG CAGCTGCACG GTGGCAGAGA TCTACCTGGG GAACCTGGCC 300
81 V F C L H K S S C T V A E I Y L G N L A 100

301 GCAGCAGACC TGATCCTGGC CTGCGGGCTG CCCTTCTGGG CCATCACCAT CTCCAACAAC 360
101 A A D L I L A C G L P F W A I T I S N N 120

361 TTCGACTGGC TCTTTGGGGA GACGCTCTGC CGCGTGGTGA ATGCCATTAT CTCCATGAAC 420
121 F D W L F G E T L C R V V N A I I S M N 140

421 CTGTACAGCA GCATCTGTTT CCTGATGCTG GTGAGCATCG ACCGCTACCT GGCCCTGGTG 480
141 L Y S S I C F L M L V S I D R Y L A L V 160

481 AAAACCATGT CCATGGGCCG GATGCGCGGC GTGCGCTGGG CCAAGCTCTA CAGCTTGGTG 540
161 K T M S M G R M R G V R W A K L Y S L V 180

541 ATCTGGGGGT GTACGCTGCT CCTGAGCTCA CCCATGCTGG TGTTCCGGAC CATGAAGGAG 600
181 I W G C T L L L S S P M L V F R T M K E 200

601 TACAGCGATG AGGGCCACAA CGTCACCGCT TGTGTCATCA GCTACCCATC CCTCATCTGG 660
201 Y S D E G H N V T A C V I S Y P S L I W 220

661 GAAGTGTTCA CCAACATGCT CCTGAATGTC GTGGGCTTCC TGCTGCCCCT GAGTGTCAATC 720
221 E V F T N M L L N V V G F L L P L S V I 240

721 ACCTTCTGCA CGATGCAGAT CATGCAGGTG CTGCGGAACA ACGAGATGCA GAAGTTCAAG 780
241 T F C T M Q I M Q V L R N N E M Q K F K 260

781 GAGATCCAGA CGGAGAGGAG GGCCACGGTG CTAGTCCTGG TTGTGCTGCT GCTATTCATC 840
261 E I Q T E R R A T V L V L V V L L L F I 280

841 ATCTGCTGGC TGCCCTTCCA GATCAGCACC TTCCTGGATA CGCTGCATCG CCTCGGCATC 900
281 I C W L P F Q I S T F L D T L H R L G I 300

```

Figure 6B

```

901 CTCTCCAGCT GCCAGGACGA GGCATCATC GATGTAATCA CACAGATCGC CTCCTTCATG 960
301 L S S C Q D E R I I D V I T Q I A S F M 320

961 GCCTACAGCA ACAGCTGCCT CAACCCACTG GTGTACGTGA TCGTGGGCAA GCGCTTCCGA 1020
321 A Y S N S C L N P L V Y V I V G K R F R 340

1021 AAGAAGTCTT GGGAGGTGTA CCAGGGAGTG TGCCAGAAAG GGGGCTGCAG GTCAGAACCC 1080
341 K K S W E V Y Q G V C Q K G G C R S E P 360

1081 ATTCAGATGG AGAACTCCAT GGGCACACTG CGGACCTCCA TCTCCGTGGA ACGCCAGATT 1140
361 I Q M E N S M G T L R T S I S V E R Q I 380

1141 CACAACTGC AGGACTGGGC AGGGAGCAGA CAGTGAGCAA ACGCCAGCAG GGCTGCTGTG 1200
381 H K L Q D W A G S R Q * 391

1201 AATTGTGTGA AGGATTGAGG GACAGTTGCT TTTCAGCATG GGCCCAGGAA TGCCAAGGAG 1260

1261 ACATCTATGC ACGACCTTGG GAAATGAGTT GATGTCTCCG GTAAACACC GGAGACTAAT 1320

1321 TCCTGCCCTG CCCAATTTTG CAGGGAGCAT GGCTGTGAGG ATGGGGTGAA CTCACGCACA 1380

1381 GCCAAGGACT CCAAAATCAC AACAGCATTA CTGTTCTTAT TTGCTGCCAC ACCTGAGCCA 1440

1441 GCCTGCTCCT TCCCAGGAGT GGAGGAGGCC TGGGGGGAGG GAGAGGAGTG ACTGAGCTTC 1500

1501 CCTCCCGTGT GTTCTCCGTC CCTGCCCCAG CAAGACAACT TAGATCTCCA GGAGAACTGC 1560

1561 CATCCAGCTT TGGTGCAATG GCTGAGTGCA CAAGTGAGTT GTTGCCCTGG GTTTCTTTAA 1620

1621 TCTATTCAGC TAGAACTTTG AAGGACAATT TCTTGCAATTA ATAAAGGTTA AGCCCTGAGG 1680

1681 GGTCCTGAT AACAACTGG AGACCAGGAT TTTATGGCTC CCCTCACTGA TGGACAAGGA 1740

1741 GGTCTGTGCC AAAGAAGAAT CCAATAAGCA CATATTGAGC ACTTGCTGTA TATGCAGTAT 1800

1801 TGAGCACTGT AGGCAAGACC CAAGAAAGAG AAGGAGCCAT CTCCATCTTG AAGGAACTCA 1860

1861 AAGACTCAAG TGGGAACGAC TGGGCACTGC CACCACCAGA AAGCTGTTCG ACGAGACGGT 1920

1921 CGAGCAGGGT GCTGTGGGTG ATATGGACAG CAGAAGGGG AGACCAAGGT TCCAGCTCAA 1980

1981 CCAATAACTA TTGCACAACC ACCTGTCCCT GCCTCAGTTC CCTTTTATGT AACATGAAGT 2040

```


Figure 6C

2041 CGTTGTGAGG GTTAAAGGCA GTAACAGGTA TAAAGTACTT AGAAAAGCAA AGGGTGCTAC 2100
 2101 GTACATGTGA GGCATCATTÀ CGCAGACGTA ACTGGGATAT GTTTACTATA AGGAAAAGAC 2160
 2161 ACTGAGGTCT AGAAATAGCT CCGTGAGGCA GAATCAGTAT TGGGAGCCGG TGGCGGTGTG 2220
 2221 AAGCACCAGT GTCTGGCACA CAGTAGGTGC TCATTGGCTC CTTCCACCT GTCATTCCCA 2280
 2281 CCACCCTGAG GCCCCAACCG CCACACACAC AGGAGCATTT GGAGAGAAGG CCATGTCTTC 2340
 2341 AAAGTCTGAT TTGTGATGAG GCAGAGGAAG ATATTTCTAA TCGGTCTTGC CCAGAGGATC 2400
 2401 ACAGTGCTGA GACCCCCCAC CACCAGCCGG TACCTGGGAA GGGGAGAGT GCAGGCCTGC 2460
 2461 TCAGGGACTG TTCCTGTCTC AGCAACCAAG GGATTGTTCC TGTCAATCAA TGGTTTATTG 2520
 2521 GAAGGTGGCC CAGTATGAGC CCTAGAAGAG TGTGAAAAGG AATGGCAATG GTGTTACCA 2580
 2581 TCGGCAGTGC CAGGGCAGCA CTCATTCACT TGATAAATGA ATATTTATTA GCTGGTTGGA 2640
 2641 GAGCTAGAAC CTGGAGAGCT AGAACCTGGA GAACTAGAAC CTGGAGGGCT AGAACCTGGA 2700
 2701 GAGGCTAGAA CCAAGAAGGG CTAGAACCTG GAGGGGCTAG AACCTAGAGA AGCTAAAACC 2760
 2761 TGAGCTAGAA GCTGGAGGAC TAGAACCTGG AGGGCTGGAA TCTGAAGGGC TAGAACCTGG 2820
 2821 AGGGCTGGAA TCTGGAGAGC TAGAACCTGG AGGGCTAGAA CCTGGAGGGC TAGAACCTAG 2880
 2881 AAGGGCTAGA ACCTGGAGGG CTGGAATCTG GAGAGCTAGA ACCTGGAGGG CTAGAACCTG 2940
 2941 GAGGGCTAGA ACCTAGAAGG GCTAGAACCT GGAGGGCTAG AACCTGGCAG GTTAGAACCT 3000
 3001 AGAAGGGCTA GAACCTGGAG AGCCAGAACC TGGAGGGCTA GAACCTGGAA GGGCTAGAAC 3060
 3061 CTGTAGAGCT AGAACATGGA GAGCTAGAAC CCGGCAGGCT AGAACCTGGC AAGCTAGAAC 3120
 3121 CTGGAGGGAA TGAACCTGGA GGGCTAGAAC CTGGAGAATG AGAAAAATTT ACATGGCAAA 3180
 3181 GAGCCCATAA ATCCTGACCA ATCCAACCTCT GAATTTTAAA GCAAAAGCGT GAAAAAAAAG 3240

Figure 6D

3241 ATTCCCTCCT TACCCCAAC CCACTCTTTT TTCCCACCAC CCACTCTCCT CTGCCTCAGT 3300
3301 AAGTATCTGG AGGAAGAAA CAGGTGAAAG AAGAAGTAA AACCATTAG TATTAGTATT 3360
3361 AGAATGAAGT CAACTGTGC CACACATGGT GAATGAAAA AAAAAAAG AGGCTGTGTT 3420
3421 TTGTCACACA GGCAGTCAT TCAGCACCAG AGCACGTGAT GGTCTGAGAC TCTCTTAGGA 3480
3481 GCAGAGCTCT GCCGCAATGG CCATGTGGGG ATCCACACCT GGTCTGAGGG GCAACTGAGT 3540
3541 CTGCGGGAGA AGAGCGGCC TATGCATGGT GTAGATGCCC TGATAAAGAA CATCTGTCCT 3600
3601 GTGAAAGACT CAATGAGCTG TTATGTTGT AAGAGGAAGC ATTTACATC CAAACGAGAA 3660
3661 AATCATGTAA ACATGTGTCT TTTCTGTAGA GCATAATAA TGGATGAGGT TTTTGCAAAA 3720
3721 AAAAAAAAA AAA 3733

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 7A

```

1  AATTCAGAGC  CACCGCGGGC  AGGCGGGCAG  TGCATCCAGA  AGCGTTTATA  TTCTGAGCGC  60

61  CAGTTCAGCT  TTCAAAAAGA  GTGCTGCCCA  TAAAAAGCCT  TCCACCCTCC  TGTCTGCTTT  120

121  AGAAGGACCC  TGAGCCCCAG  GCGCCAGCCA  CAGGACTCTG  CTGCAGAGGG  GGGTTGTGTA  180

181  CAGATAGTAG  GCTTTACGCC  TAGCTTCGAA  ATGGATAACG  TCCTCCCGGT  GGACTCAGAC  240
1      M  D  N  V      L  P  V      D  S  D      10

241  CTCTCCCCAA  ACATCTCCAC  TAACACCTCG  GAACCCAATC  AGTTCGTGCA  ACCAGCCTGG  300
11  L  S  P  N      I  S  T      N  T  S      E  P  N  Q      F  V  Q      P  A  W      30

301  CAAATTGTCC  TTTGGGCAGC  TGCCTACACG  GTCATTGTGG  TGACCTCTGT  GGTGGGCAAC  360
31  Q  I  V  L      W  A  A      A  Y  T      V  I  V  V      T  S  V      V  G  N      50

361  GTGGTAGTGA  TGTGGATCAT  CTTAGCCCAC  AAAAGAATGA  GGACAGTGAC  GAACTATTTT  420
51  V  V  V  M      W  I  I      L  A  H      K  R  M  R      T  V  T      N  Y  F      70

421  CTGGTGAACC  TGGCCTTCGC  GGAGGCCTCC  ATGGCTGCAT  TCAATACAGT  GGTGAACTTC  480
71  L  V  N  L      A  F  A      E  A  S      M  A  A  F      N  T  V      V  N  F      90

481  ACCTATGCTG  TCCACAACGA  ATGGTACTAC  GGCCTGTCTT  ACTGCAAGTT  CCACAACCTC  540
91  T  Y  A  V      H  N  E      W  Y  Y      G  L  F  Y      C  K  F      H  N  F      110

541  TTTCCCATCG  CCGCTGTCTT  CGCCAGTATC  TACTCCATGA  CGGCTGTGGC  CTTTGATAGG  600
111  F  P  I  A      A  V  F      A  S  I      Y  S  M  T      A  V  A      F  D  R      130

601  TACATGGCCA  TCATACATCC  CCTCCAGCCC  CGGCTGTCAG  CCACAGCCAC  CAAAGTGGTC  660
131  Y  M  A  I      I  H  P      L  Q  P      R  L  S  A      T  A  T      K  V  V      150

661  ATCTGTGTCA  TCTGGGTCCT  GGCTCTCCTG  CTGGCCTTCC  CCCAGGGCTA  CTA CTCAACC  720
151  I  C  V  I      W  V  L      A  L  L      L  A  F  P      Q  G  Y      Y  S  T      170

721  ACAGAGACCA  TGCCAGCAG  AGTCGTGTGC  ATGATCGAAT  GGCCAGAGCA  TCCGAACAAG  780
171  T  E  T  M      P  S  R      V  V  C      M  I  E  W      P  E  H      P  N  K      190

781  ATTTATGAGA  AAGTGTAACA  CATCTGTGTG  ACTGTGCTGA  TCTACTTCCT  CCCCTGCTG  840
191  I  Y  E  K      V  Y  H      I  C  V      T  V  L  I      Y  F  L      P  L  L      210

841  GTGATTGGCT  ATGCATACAC  CGTAGTGGGA  ATCACACTAT  GGGCCAGTGA  GATCCCCGGG  900
211  V  I  G  Y      A  Y  T      V  V  G      I  T  L  W      A  S  E      I  P  G      230

901  GACTCCTCTG  ACCGCTACCA  CGAGCAAGTC  TCTGCCAAGC  GCAAGGTGGT  CAAAATGATG  960
231  D  S  S  D      R  Y  H      E  Q  V      S  A  K  R      K  V  V      K  M  M      250

```

Figure 7B

```

961 ATTGTCGTGG TGTGCACCTT CGCCATCTGC TGGCTGCCCT TCCACATCTT CTTCCTCCTG 1020
251 I V V V C T F A I C W L P F H I F F L L 270

1021 CCCTACATCA ACCCAGATCT CTACCTGAAG AAGTTTATCC AGCAGGTCTA CCTGGCCATC 1080
271 P Y I N P D L Y L K K F I Q Q V Y L A I 290

1081 ATGTGGCTGG CCATGAGCTC CACCATGTAC AACCCCATCA TCTACTGCTG CCTCAATGAC 1140
291 M W L A M S S T M Y N P I I Y C C L N D 310

1141 AGGTTCGGTC TGGGCTTCAA GCATGCCTTC CGGTGCTGCC CTTTCATCAG CGCCGGCGAC 1200
311 R F R L G F K H A F R C C P F I S A G D 330

1201 TATGAGGGGC TGGAAATGAA ATCCACCCGG TATCTCCAGA CCCAGGGCAG TGTGTACAAA 1260
331 Y E G L E M K S T R Y L Q T Q G S V Y K 350

1261 GTCAGCCGCC TGGAGACCAC CATCTCCACA GTGGTGGGGG CCCACGAGGA GGAGCCAGAG 1320
351 V S R L E T T I S T V V G A H E E E P E 370

1321 GACGGCCCCA AGGCCACACC CTCGTCCCTG GACCTGACCT CCAACTGCTC TTCACGAAGT 1380
371 D G P K A T P S S L D L T S N C S S R S 390

1381 GACTCCAAGA CCATGACAGA GAGCTTCAGC TTCTCCTCCA ATGTGCTCTC CTAGGCCACA 1440
391 D S K T M T E S F S F S S N V L S * 407

1441 GGGCCTTTGG CAGGTGCAGC CCCCCTGCC TTTGACCTGC CTCCCTTCAT GCATGGAAAT 1500

1501 TCCCTTCATC TGGAACCATC AGAAACACCC TCACACTGGG ACTTGCAAAA AGGGTCAGTA 1560

1561 TGGGTAGGG AAAACATTCC ATCCTTGAGT CAAAAAATCT CAATTCTTCC CTATCTTTGC 1620

1621 CACCCTCATG CTGTGTGACT CAAACCAAAT CACTGAACTT TGCTGAGCCT GTAAATATAA 1680

1681 AGGTCGGACC AGCTTTTCCT CAAGAGCCCA ATGCATTCCA TTTCTGGAAG TGACTTTGGC 1740

1741 TGCATGCGAG TGCTCATTTT AGGATG 1766

```

Figure 8A

```

1 AATTCAGAGC CACCGCGGGC AGGCGGGCAG TGCATCCAGA AGCGTTTATA TTCTGAGCGC 60

61 CAGTTCAGCT TTCAAAAAGA GTGCTGCCCCA TAAAAAGCCT TCCACCCTCC TGTCTGCTTT 120

121 AGAAGGACCC TGAGCCCCAG GCGCCAGCCA CAGGACTCTG CTGCAGAGGG GGGTTGTGTA 180

181 CAGATAGTAG GCTTTACGCC TAGCTTCGAA ATGGATAACG TCCTCCCGGT GGACTCAGAC 240
1 M D N V L P V D S D 10

241 CTCTCCCCAA ACATCTCCAC TAACACCTCG GAACCCAATC AGTTCGTGCA ACCAGCCTGG 300
11 L S P N I S T N T S E P N Q F V Q P A W 30

301 CAAATTGTCC TTGGGCAGC TGCCTACACG GTCATTGTGG TGACCTCTGT GGTGGGCAAC 360
31 Q I V L W A A A Y T V I V V T S V V G N 50

361 GTGGTAGTGA TGTGGATCAT CTTAGCCCAC AAAAGAATGA GGACAGTGAC GAACTATTTT 420
51 V V V M W I I L A H K R M R T V T N Y F 70

421 CTGGTGAACC TGGCCTTCGC GGAGGCCTCC ATGGCTGCAT TCAATACAGT GGTGAACTTC 480
71 L V N L A F A E A S M A A F N T V V N F 90

481 ACCTATGCTG TCCACAACGA ATGGTACTAC GGCCTGTCTT ACTGCAAGTT CCACAACCTC 540
91 T Y A V H N E W Y Y G L F Y C K F H N F 110

541 TTCCCCATCG CCGCTGTCTT CGCCAGTATC TACTCCATGA CGGCTGTGGC CTTTGATAGG 600
111 F P I A A V F A S I Y S M T A V A F D R 130

601 TACATGGCCA TCATACATCC CCTCCAGCCC CGGCTGTCAG CCACAGCCAC CAAAGTGGTC 660
131 Y M A I I H P L Q P R L S A T A T K V V 150

661 ATCTGTGTCA TCTGGGTCCT GGCTCTCCTG CTGGCCTTCC CCCAGGGCTA CTACTCAACC 720
151 I C V I W V L A L L L A F P Q G Y Y S T 170

721 ACAGAGACCA TGCCCAGCAG AGTCGTGTGC ATGATCGAAT GGCCAGAGCA TCCGAACAAG 780
171 T E T M P S R V V C M I E W P E H P N K 190

781 ATTTATGAGA AAGTGTAACA CATCTGTGTG ACTGTGCTGA TCTACTTCCT CCCCTGCTG 840
191 I Y E K V Y H I C V T V L I Y F L P L L 210

841 GTGATTGGCT ATGCATACAC CGTAGTGGGA ATCACACTAT GGGCCAGTGA GATCCCCGGG 900
211 V I G Y A Y T V V G I T L W A S E I P G 230

901 GACTCCTCTG ACCGCTACCA CGAGCAAGTC TCTGCCAAGC GCAAGGTGGT CAAAATGATG 960
231 D S S D R Y H E Q V S A K R K V V K M M 250

```

Figure 8B

```

961 ATTGTCGTGG TGTGCACCTT CGCCATCTGC TGGCTGCCCT TCCACATCTT CTTCCCTCCTG 1020
251 I V V V C T F A I C W L P F H I F F L L 270

1021 CCCTACATCA ACCCAGATCT CTACCTGAAG AAGTTTATCC AGCAGGTCTA CCTGGCCATC 1080
271 P Y I N P D L Y L K K F I Q Q V Y L A I 290

1081 ATGTGGCTGG CCATGAGCTC CACCATGTAC AACCCCATCA TCTACTGCTG CCTCAATGAC 1140
291 M W L A M S S T M Y N P I I Y C C L N D 310

1141 AGGTTCCGTC TGGGCTTCAA GCATGCCTTC CGGTGCTGCC CCTTCATCAG CGCCGGCGAC 1200
311 R F R L G F K H A F R C C P F I S A G D 330

1201 TATGAGGGGC TGGAAATGAA ATCCACCCGG TATCTCCAGA CCCAGGGCAG TGTGTACAAA 1260
331 Y E G L E M K S T R Y L Q T Q G S V Y K 350

1261 GTCAGCCGCC TGGAGACCAC CATCTCCACA GTGGTGGGGG CCCACGAGGA GGAGCCAGAG 1320
351 V S R L E T T I S T V V G A H E E E P E 370

1321 GACGGCCCCA AGGCCACACC CTCGTCCCTG GACCTGACCT CCAACTGCTC TTCACGAAGT 1380
371 D G P K A T P S S L D L T S N C S S R S 390

1381 GACTCCAAGA CCATGACAGA GAGCTTCAGC TTCTCCTCCA ATGTGCTCTC CTAGGCCACA 1440
391 D S K T M T E S F S F S S N V L S * 407

1441 GGGCCTTTGG CAGGTGCAGC CCCCACTGCC TTTGACCTGC CTCCCTTCAT GCATGGAAAT 1500

1501 TCCCTTCATC TGGAACCATC AGAAACACCC TCACACTGGG ACTTGCAAAA AGGGTCAGTA 1560

1561 TGGGTTAGGG AAAACATTCC ATCCTTGAGT CAAAAAATCT CAATTCTTCC CTATCTTTGC 1620

1621 CACCCTCATG CTGTGTGACT CAAACCAAAT CACTGAACTT TGCTGAGCCT GTAAAATAAA 1680

1681 AGGTCGGACC AGCTTTTCCT CAAGAGCCCA ATGCATTCCA TTTCTGGAAG TGACTTTGGC 1740

1741 TGCATGCGAG TGCTCATTTT AGGATG 1766

```

Figure 9A

```

1  AATTCAGAGC CACCGCGGGC AGGCGGGCAG TGCATCCAGA AGCGTTTATA TTCTGAGCGC 60

61  CAGTTCAGCT TTCAAAAAGA GTGCTGCCCA TAAAAAGCCT TCCACCCTCC TGTCTGCTTT 120

121 AGAAGGACCC TGAGCCCCAG GCGCCAGCCA CAGGACTCTG CTGCAGAGGG GGGTTGTGTA 180

181 CAGATAGTAG GCTTTACGCC TAGCTTCGAA ATGGATAACG TCCTCCCGGT GGACTCAGAC 240
    1          M D N V L P V D S D 10

241 CTCTCCCCAA ACATCTCCAC TAACACCTCG GAACCCAATC AGTTCGTGCA ACCAGCCTGG 300
    11 L S P N I S T N T S E P N Q F V Q P A W 30

301 CAAATTGTCC TTTGGGCAGC TGCCTACACG GTCATGTGGG TGACCTCTGT GGTGGGCAAC 360
    31 Q I V L W A A A Y T V I V V T S V V G N 50

361 GTGGTAGTGA TGTGGATCAT CTTAGCCCAC AAAAGAATGA GGACAGTGAC GAACTATTTT 420
    51 V V V M W I I L A H K R M R T V T N Y F 70

421 CTGGTGAACC TGGCCTTCGC GGAGGCCTCC ATGGCTGCAT TCAATACAGT GGTGAAC TTC 480
    71 L V N L A F A E A S M A A F N T V V N F 90

481 ACCTATGCTG TCCACAACGA ATGGTACTAC GGCCTGTCTC ACTGCAAGTT CCACAAC TTC 540
    91 T Y A V H N E W Y Y G L F Y C K F H N F 110

541 TTTCCCATCG CCGCTGTCTT CGCCAGTATC TACTCCATGA CGGCTGTGGC CTTTGATAGG 600
    111 F P I A A V F A S I Y S M T A V A F D R 130

601 TACATGGCCA TCATACATCC CCTCCAGCCC CGGCTGTCAG CCACAGCCAC CAAAGTGGTC 660
    131 Y M A I I H P L Q P R L S A T A T K V V 150

661 ATCTGTGTCA TATGGGTCCT GGCTCTCCTG CTGGCCTTCC CCCAGGGCTA CTA CTCAACC 720
    151 I C V I W V L A L L L A F P Q G Y Y S T 170

721 ACAGAGACCA TGCCCAGCAG AGTCGTGTGC ATGATCGAAT GGCCAGAGCA TCCGAACAAG 780
    171 T E T M P S R V V C M I E W P E H P N K 190

781 ATTTATGAGA AAGTGACCA CATCTGTGTG ACTGTGCTGA TCTACTTCCT CCCCCTGCTG 840
    191 I Y E K V Y H I C V T V L I Y F L P L L 210

841 GTGATTGGCT ATGCATACAC CGTAGTGGGA ATCACACTAT GGGCCAGTGA GATCCCCGGG 900
    211 V I G Y A Y T V V G I T L W A S E I P G 230

901 GACTCCTCTG ACCGCTACCA CGAGCAAGTC TCTGCCAAGC GCAAGGTGGT CAAAATGATG 960
    231 D S S D R Y H E Q V S A K R K V V K M M 250

```

Figure 9B

```

961 ATTGTCGTGG TGTGCACCTT CGCCATCTGC TGGCTGCCCT TCCACATCTT CTCCTCCTG 1020
251 I V V V C T F A I C W L P F H I F F L L 270

1021 CCCTACATCA ACCCAGATCT CTACCTGAAG AAGTTTATCC AGCAGGTCTA CCTGGCCATC 1080
271 P Y I N P D L Y L K K F I Q Q V Y L A I 290

1081 ATGTGGCTGG CCATGAGCTC CACCATGTAC AACCCCATCA TCTACTGCTG CCTCAATGAC 1140
291 M W L A M S S T M Y N P I I Y C C L N D 310

1141 AGGTTCGTC TGGGCTTCAA GCATGCCTTC CGGTGCTGCC CCTTCATCAG CGCCGGCGAC 1200
311 R F R L G F K H A F R C C P F I S A G D 330

1201 TATGAGGGGC TGGAAATGAA ATCCACCCGG TATCTCCAGA CCCAGGGCAG TGTGTACAAA 1260
331 Y E G L E M K S T R Y L Q T Q G S V Y K 350

1261 GTCAGCCGCC TGGAGACCAC CATCTCCACA GTGGTGGGGG CCCACGAGGA GGAGCCAGAG 1320
351 V S R L E T T I S T V V G A H E E E P E 370

1321 GACGGCCCCA AGGCCACACC CTCGTCCCTG GACCTGACCT CCAACTGCTC TTCACGAAGT 1380
371 D G P K A T P S S L D L T S N C S S R S 390

1381 GACTCCAAGA CCATGACAGA GAGCTTCAGC TTCTCCTCCA ATGTGCTCTC CTAGGCCACA 1440
391 D S K T M T E S F S F S S N V L S * 407

1441 GGGCCTTTGG CAGGTGCAGC CCCCCTGCC TTTGACCTGC CTCCCTTCAT GCATGGAAAT 1500

1501 TCCCTTCATC TGGAACCATC AGAAACACCC TCACACTGGG ACTTGCAAAA AGGGTCAGTA 1560

1561 TGGGTTAGGG AAAACATTCC ATCCTTGAGT CAAAAAATCT CAATTCTTCC CTATCTTTGC 1620

1621 CACCCTCATG CTGTGTGACT CAAACCAAAT CACTGAACTT TGCTGAGCCT GTAAAATAAA 1680

1681 AGGTCGGACC AGCTTTTCCT CAAGAGCCCA ATGCATTCCA TTTCTGGAAG TGACTTTGGC 1740

1741 TGCATGCGAG TGCTCATTTT AGGATG 1766

```


Figure 10A

1 AATTCAGAGC CACCGCGGGC AGGCGGGCAG TGCATCCAGA AGCGTTTATA TTCTGAGCGC 60
 61 CAGTTCAGCT TTCAAAAAGA GTGCTGCCCA TAAAAAGCCT TCCACCCTCC TGTCTGCTTT 120
 121 AGAAGGACCC TGAGCCCCAG GCGCCAGCCA CAGGACTCTG CTGCAGAGGG GGGTTGTGTA 180
 181 CAGATAGTAG GCTTTACGCC TAGCTTCGAA ATGGATAACG TCCTCCCGGT GGACTCAGAC 240
 1 M D N V L P V D S D 10
 241 CTCTCCCCAA ACATCTCCAC TAACACCTCG GAACCCAATC AGTTCGTGCA ACCAGCCTGG 300
 11 L S P N I S T N T S E P N Q F V Q P A W 30
 301 CAAATTGTCC TTTGGGCAGC TGCCTACACG GTCATTGTGG TGACCTCTGT GGTGGGCAAC 360
 31 Q I V L W A A A Y T V I V V T S V V G N 50
 361 GTGGTAGTGA TGTGGATCAT CTTAGCCCAC AAAAGAATGA GGACAGTGAC GAACTATTTT 420
 51 V V V M W I I L A H K R M R T V T N Y F 70
 421 CTGGTGAACC TGGCCTTCGC GGAGGCCTCC ATGGCTGCAT TCAATACAGT GGTGAACTTC 480
 71 L V N L A F A E A S M A A F N T V V N F 90
 481 ACCTATGCTG TCCACAACGA ATGGTACTAC GGCCTGTCTT ACTGCAAGTT CCACAACCTC 540
 91 T Y A V H N E W Y Y G L F Y C K F H N F 110
 541 TTTCCCATCG CCGCTGTCTT CGCCAGTATC TACTCCATGA CGGCTGTGGC CTTTGATAGG 600
 111 F P I A A V F A S I Y S M T A V A F D R 130
 601 TACATGGCCA TCATACATCC CCTCCAGCCC CGGCTGTCAG CCACAGCCAC CAAAGTGGTC 660
 131 Y M A I I H P L Q P R L S A T A T K V V 150
 661 ATCTGTGTCA TCTGGGTCCT GGCTCTCCTG CTGGCCTTCC CCCAGGGCTA CTA CTCAACC 720
 151 I C V I W V L A L L L A F P Q G Y Y S T 170
 721 ACAGAGACCA TGCCCAGCAG AGTCGTGTGC ATGATCGAAT GGCCAGAGCA TCCGAACAAG 780
 171 T E T M P S R V V C M I E W P E H P N K 190
 781 ATTTATGAGA AAGTGTAACA CATCTGTGTG ACTGTGCTGA TCTACTTCCT CCCCTGCTG 840
 191 I Y E K V Y H I C V T V L I Y F L P L L 210
 841 GTGATTGGCT ATGCATACAC CGTAGTGGGA ATCACACTAT GGGCCAGTGA GATCCCCGGG 900
 211 V I G Y A Y T V V G I T L W A S E I P G 230
 901 GACTCCTCTG ACCGCTACCA CGAGCAAGTC TCTGCCAAGC GCAAGGTGGT CAAAATGATG 960
 231 D S S D R Y H E Q V S A K R K V V K M M 250

Figure 10B

```

961 ATTGTCGTGG TGTGCACCTT CGCCATCTGC TGGCTGCCCT TCCACATCTT CTTCTCCTG 1020
251 I V V V C T F A I C W L P F H I F F L L 270

1021 CCCTACATCA ACCCAGATCT CTACCTGAAG AAGTTTATCC AGCAGGTCTA CCTGGCCATC 1080
271 P Y I N P D L Y L K K F I Q Q V Y L A I 290

1081 ATGTGGCTGG CCATGAGCTC CACCATGTAC AACCCCATCA TCTACTGCTG CCTCAATGAC 1140
291 M W L A M S S T M Y N P I I Y C C L N D 310

1141 AGGTTCCGTC TGGGCTTCAA GCATGCCTTC CGGTGCTGCC CCTTCATCAG CGCCGGCGAC 1200
311 R F R L G F K H A F R C C P F I S A G D 330

1201 TATGAGGGGC TGGAAATGAA ATCCACCCGG TATCTCCAGA CCCAGGGCAG TGTGTACAAA 1260
331 Y E G L E M K S T R Y L Q T Q G S V Y K 350

1261 GTCAGCCGCC TGGAGACCA C CATCTCCACA GTGGTGGGGG CCCACGAGGA GGAGCCAGAG 1320
351 V S R L E T T I S T V V G A H E E E P E 370

1321 GACGGCCCCA AGGCCACACC CTCATCCCTG GACCTGACCT CCAACTGCTC TTCACGAAGT 1380
371 D G P K A T P S S L D L T S N C S S R S 390

1381 GACTCCAAGA CCATGACAGA GAGCTTCAGC TTCTCCTCCA ATGTGCTCTC CTAGGCCACA 1440
391 D S K T M T E S F S F S S N V L S * 407

1441 GGGCCTTTGG CAGGTGCAGC CCCCCTGACC TTTGACCTGC CTCCCTTCAT GCATGGAAAT 1500

1501 TCCCTTCATC TGGAACCATC AGAAACACCC TCACACTGGG ACTTGCAAAA AGGGTCAGTA 1560

1561 TGGGTTAGGG AAAACATTCC ATCCTTGAGT CAAAAAATCT CAATTCTTCC CTATCTTTGC 1620

1621 CACCCTCATG CTGTGTGACT CAAACCAAAT CACTGAACTT TGCTGAGCCT GTAAAATAAA 1680

1681 AGGTCGGACC AGCTTTTCCT CAAGAGCCCA ATGCATTCCA TTTCTGGAAG TGACTTTGGC 1740

1741 TGCATGCGAG TGCTCATTTT AGGATG 1766

```

Figure 11A

```

1 AGTCTGCACT GGAGCTGCCT GGTGACCAGA AGTTTGGAGT CCGCTGACGT CGCCGCCAG 60

61 ATGGCCTCCA GGCTGACCCT GCTGACCCTC CTGCTGCTGC TGCTGGCTGG GGATAGAGCC 120
1 M A S R L T L L T L L L L L L A G D R A 20

121 TCCTCAAATC CAAATGCTAC CAGCTCCAGC TCCCAGGATC CAGAGAGTTT GCAAGACAGA 180
21 S S N P N A T S S S S Q D P E S L Q D R 40

181 GGCGAAGGGA AGGTGCAAC AACAGTTATC TCCAAGATGC TATTCGTTGA ACCCATCCTG 240
41 G E G K V A T T V I S K M L F V E P I L 60

241 GAGGTTTCCA GCTTGCCGAC AACCAACTCA ACAACCAATT CAGCCACCAA AATAACAGCT 300
61 E V S S L P T T N S T T N S A T K I T A 80

301 AATACCACTG ATGAACCCAC CACACAACCC ACCACAGAGC CCACCACCCA ACCCACCATC 360
81 N T T D E P T T Q P T T E P T T Q P T I 100

361 CAACCCACCC AACCAACTAC CCAGCTCCCA ACAGATTCTC CTACCCAGCC CACTACTGGG 420
101 Q P T Q P T T Q L P T D S P T Q P T T G 120

421 TCCTTCTGCC CAGGACCTGT TACTCTCTGC TCTGACTTGG AGAGTCATTC AACAGAGGCC 480
121 S F C P G P V T L C S D L E S H S T E A 140

481 GTGTTGGGGG ATGCTTTGGT AGATTTCTCC CTGAAGCTCT ACCACGCCTT CTCAGCAATG 540
141 V L G D A L V D F S L K L Y H A F S A M 160

541 AAGAAGGTGG AGACCAACAT GGCCTTTTCC CCATTGAGCA TCGCCAGCCT CTTTACCCAG 600
161 K K V E T N M A F S P F S I A S L L T Q 180

601 GTCCTGCTCG GGGCTGGGCA GAACACCAA ACAAACCTGG AGAGCATCCT CTCTTACCCC 660
181 V L L G A G Q N T K T N L E S I L S Y P 200

661 AAGGACTTCA CCTGTGTCCA CCAGGCCCTG AAGGGCTTCA CGACCAAAGG TGTACCTCA 720
201 K D F T C V H Q A L K G F T T K G V T S 220

721 GTCTCTCAGA TCTTCCACAG CCCAGACCTG GCCATAAGGG ACACCTTTGT GAATGCCTCT 780
221 V S Q I F H S P D L A I R D T F V N A S 240

781 CGGACCCTGT ACAGCAGCAG CCCAGAGTC CTAAGCAACA ACAGTGACGC CAACTTGGAG 840
241 R T L Y S S S P R V L S N N S D A N L E 260

841 CTCATCAACA CCTGGGTGGC CAAGAACACC AACAACAAGA TCAGCCGGCT GCTAGACAGT 900
261 L I N T W V A K N T N N K I S R L L D S 280

901 CTGCCCTCCG ATACCCGCCT TGTCTCTCTC AATGCTATCT ACCTGAGTGC CAAGTGAAG 960
281 L P S D T R L V L L N A I Y L S A K W K 300

```

Figure 11B

```

961 ACAACATTTG ATCCCAAGAA AACCAGAATG GAACCCTTTC ACTTCAAAAA CTCAGTTATA 1020
301 T T F D P K K T R M E P F H F K N S V I 320

1021 AAAGTGCCCA TGATGAATAG CAAGAAGTAC CCTGTGGCCC ATTCATTGA CCAAACCTTG 1080
321 K V P M M N S K K Y P V A H F I D Q T L 340

1081 AAAGCCAAGG TGGGGCAGCT GCAGCTCTCC CACAATCTGA GTTTGGTGAT CCTGGTACCC 1140
341 K A K V G Q L Q L S H N L S L V I L V P 360

1141 CAGAACCTGA AACATCGTCT TGAAGACATG GAACAGGCTC TCAGCCCTTC TGTTTTCAAG 1200
361 Q N L K H R L E D M E Q A L S P S V F K 380

1201 GCCATCATGG AGAAACTGGA GATGTCCAAG TTCCAGCCCA CTCTCCTAAC ACTACCCCGC 1260
381 A I M E K L E M S K F Q P T L L T L P R 400

1261 ATCAAAGTGA CGACCAGCCA GGATATGCTC TCAATCATGG AGAAATTGGA ATTCTTCGAT 1320
401 I K V T T S Q D M L S I M E K L E F F D 420

1321 TTTTCTTATG ACCTTAACCT GTGTGGGCTG ACAGAGGACC CAGATCTTCA GGTTCCTGCG 1380
421 F S Y D L N L C G L T E D P D L Q V S A 440

1381 ATGCAGCACC AGACAGTGCT GGAAGTACA GAGACTGGGG TGGAGGCGGC TGCAGCCTCC 1440
441 M Q H Q T V L E L T E T G V E A A A A S 460

1441 GCCATCTCTG TGGCCCGCAC CCTGCTGGTC TTTGAAGTGC AGCAGCCCTT CCTCTTCGTG 1500
461 A I S V A R T L L V F E V Q Q P F L F V 480

1501 CTCTGGGACC AGCAGCACAA GTTCCTGTCT TTCATGGGGC GAGTATATGA CCCCAGGGCC 1560
481 L W D Q Q H K F P V F M G R V Y D P R A 500

1561 TGAGACCTGC AGGATCAGGT TAGGGCGAGC GCTACCTCTC CAGCCTCAGC TCTCAGTTGC 1620

1621 AGCCCTGCTG CTGCCTGCCT GGAAGTGCCC CTGCCACCTC CTGCCTCAGG TGTCCGCTAT 1680

1681 CCACCAAAAG GGCTCCTGAG GGTCTGGGCA AGGGACCTGC TTCTATTAGC CCTTCTCCAT 1740

1741 GGCCCTGCCA TGCTCTCCAA ACCACTTTTT GCAGCTTTCT CTAGTTCAAG TTCACCAGAC 1800

1801 TCTATAAATA AAACCTGACA GACCAT 1826

```

Figure 12A

```

1 AGTCTGCACT GGAGCTGCCT GGTGACCAGA AGTTTGGAGT CCGCTGACGT CGCCGCCCAG 60

61 ATGGCCTCCA GGCTGACCCT GCTGACCCTC CTGCTGCTGC TGCTGGCTGG GGATAGAGCC 120
1 M A S R L T L L T L L L L L A G D R A 20

121 TCCTCAAATC CAAATGCTAC CAGCTCCAGC TCCCAGGATC CAGAGAGTTT GCAAGACAGA 180
21 S S N P N A T S S S S Q D P E S L Q D R 40

181 GCGAAGGGA AGGTCGCAAC AACAGTTATC TCCAAGATGC TATTCGTTGA ACCCATCCTG 240
41 G E G K V A T T V I S K M L F V E P I L 60

241 GAGGTTTCCA GCTTGCCGAC AACCAACTCA ACAACCAATT CAGCCACCAA AATAACAGCT 300
61 E V S S L P T T N S T T N S A T K I T A 80

301 AATACCACTG ATGAACCCAC CACACAACCC ACCACAGAGC CCACCACCCA ACCCACCATC 360
81 N T T D E P T T Q P T T E P T T Q P T I 100

361 CAACCCACCC AACCAACTAC CCAGCTCCCA ACAGATTCTC CTACCCAGCC CACTACTGGG 420
101 Q P T Q P T T Q L P T D S P T Q P T T G 120

421 TCCTTCTGCC CAGGACCTGT TACTCTCTGC TCTGACTTGG AGAGTCATTC AACAGAGGCC 480
121 S F C P G P V T L C S D L E S H S T E A 140

481 GTGTTGGGGG ATGCTTTGGT AGATTTCTCC CTGAAGCTCT ACCACGCCTT CTCAGCAATG 540
141 V L G D A L V D F S L K L Y H A F S A M 160

541 AAGAAGGTGG AGACCAACAT GGCCTTTTCC CCATTACAGC TCGCCAGCCT CTTATCCAG 600
161 K K V E T N M A F S P F S I A S L L T Q 180

601 GTCCTGCTCG GGCTGGGCA GAACACCAAA ACAACCTGG AGAGCATCCT CTCTTACCCC 660
181 V L L G A G Q N T K T N L E S I L S Y P 200

661 AAGGACTTCA CCTGTGTCCA CCAGGCCCTG AAGGGCTTCA CGACCAAAGG TGTCACCTCA 720
201 K D F T C V H Q A L K G F T T K G V T S 220

721 GTCTCTCAGA TCTTCCACAG CCCAGACCTG GCCATAAGGG ACACCTTTGT GAATGCCTCT 780
221 V S Q I F H S P D L A I R D T F V N A S 240

781 CGGACCCTGT ACAGCAGCAG CCCAGAGTCT CTAAGCAACA ACAGTGACGC CAACTTGGAG 840
241 R T L Y S S S P R V L S N N S D A N L E 260

841 CTCATCAACA CCTGGGTGGC CAAGAACACC AACAACAAGA TCAGCCGGCT GCTAGACAGT 900
261 L I N T W V A K N T N N K I S R L L D S 280

901 CTGCCCTCCG ATACCCGCCT TGTCTCCTC AATGCTATCT ACCTGAGTGC CAAGTGAAG 960
281 L P S D T R L V L L N A I Y L S A K W K 300

```

Figure 12B

961 ACAACATTG· ATCCCAAGAA· AACCAGAATG· GAACCCTTTC· ACTTCAAAAA· CTCAGTTATA· 1020
 301 T T F D P K K T R M E P F H F K N S V I 320

1021 AAAGTGCCCA· TGATGAATAG· CAAGAAGTAC· CCTGTGGCCC· ATTCATTGA· CCAAACCTTG· 1080
 321 K V P M M N S K K Y P V A H F I D Q T L 340

1081 AAAGCCAAGG· TGGGGCAGCT· GCAGCTCTCC· CACAATCTGA· GTTGGGTGAT· CCTGGTACCC· 1140
 341 K A K V G Q L Q L S H N L S L V I L V P 360

1141 CAGAACCTGA· AACATCGTCT· TGAAGACATG· GAACAGGCTC· TCAGCCCTTC· TGTTCCTCAAG· 1200
 361 Q N L K H R L E D M E Q A L S P S V F K 380

1201 GCCATCATGG· AGAAACTGGA· GATGTCCAAG· TTCCAGCCCA· CTCTCCTAAC· ACTACCCCGC· 1260
 381 A I M E K L E M S K F Q P T L L T L P R 400

1261 ATCAAAGTGA· CGACCAGTCA· GGATATGCTC· TCAATCATGG· AGAAATTGGA· ATTCTTCGAT· 1320
 401 I K V T T S Q D M L S I M E K L E F F D 420

1321 TTTCTTATG· ACCTTAACCT· GTGTGGGCTG· ACAGAGGACC· CAGATCTTCA· GGTTTCTGCG· 1380
 421 F S Y D L N L C G L T E D P D L Q V S A 440

1381 ATGCAGCACC· AGACAGTGCT· GGAAGTACA· GAGACTGGGG· TGGAGGCGGC· TGCAGCCTCC· 1440
 441 M Q H Q T V L E L T E T G V E A A A A S 460

1441 GCCATCTCTG· TGGCCCGCAC· CCTGCTGGTC· TTTGAAGTGC· AGCAGCCCTT· CCTCTTCGTG· 1500
 461 A I S V A R T L L V F E V Q Q P F L F V 480

1501 CTCTGGGACC· AGCAGCACAA· GTTCCCTGTC· TTCATGGGGC· GAGTATATGA· CCCCAGGGCC· 1560
 481 L W D Q Q H K F P V F M G R V Y D P R A 500

1561 TGAGACCTGC· AGGATCAGGT· TAGGGCGAGC· GCTACCTCTC· CAGCCTCAGC· TCTCAGTTGC· 1620

1621 AGCCCTGCTG· CTGCCTGCCT· GGAAGTGCCC· CTGCCACCTC· CTGCCTCAGG· TGTCCGCTAT· 1680

1681 CCACCAAAAG· GGCTCCTGAG· GGTCTGGGCA· AGGGACCTGC· TTCTATTAGC· CCTTCTCCAT· 1740

1741 GGCCCTGCCA· TGCTCTCCAA· ACCACTTTTT· GCAGCTTTCT· CTAGTTCAAG· TTCACCAGAC· 1800

1801 TCTATAAATA· AACCTGACA· GACCAT 1826

Figure 13A

```

1 AGTCTGCACT GGAGCTGCCT GGTGACCAGA AGTTTGGAGT CCGCTGACGT CGCCGCCCAG 60

61 ATGGCCTCCA GGCTGACCCT GCTGACCCTC CTGCTGCTGC TGCTGGCTGG GGATAGAGCC 120
1 M A S R L T L L T L L L L L L A G D R A 20

121 TCCTCAAATC CAAATGCTAC CAGCTCCAGC TCCCAGGATC CAGAGAGTTT GCAAGACAGA 180
21 S S N P N A T S S S S Q D P E S L Q D R 40

181 GGCGAAGGGA AGGTGCAAC AACAGTTATC TCCAAGATGC TATTCGCTGA ACCCATCCTG 240
41 G E G K V A T T V I S K M L F A E P I L 60

241 GAGGTTTCCA GCTTGCCGAC AACCAACTCA ACAACCAATT CAGCCACCAA AATAACAGCT 300
61 E V S S L P T T N S T T N S A T K I T A 80

301 AATACCACTG ATGAACCCAC CACACAACCC ACCACAGAGC CCACCACCCA ACCCACCATC 360
81 N T T D E P T T Q P T T E P T T Q P T I 100

361 CAACCCACCC AACCAACTAC CCAGCTCCCA ACAGATTCTC CTACCCAGCC CACTACTGGG 420
101 Q P T Q P T T Q L P T D S P T Q P T T G 120

421 TCCTTCTGCC CAGGACCTGT TACTCTCTGC TCTGACTTGG AGAGTCATTC AACAGAGGCC 480
121 S F C P G P V T L C S D L E S H S T E A 140

481 GTGTTGGGGG ATGCTTTGGT AGATTTCTCC CTGAAGCTCT ACCACGCCTT CTCAGCAATG 540
141 V L G D A L V D F S L K L Y H A F S A M 160

541 AAGAAGGTGG AGACCAACAT GGCTTTTCC CCATTGAGCA TCGCCAGCCT CTTTACCCAG 600
161 K K V E T N M A F S P F S I A S L L T Q 180

601 GTCCTGCTCG GGCTGGGCA GAACACCAA ACAACCTGG AGAGCATCCT CTCTTACCCC 660
181 V L L G A G Q N T K T N L E S I L S Y P 200

661 AAGGACTTCA CCTGTGTCCA CCAGGCCCTG AAGGGCTTCA CGACCAAAGG TGTACCTCA 720
201 K D F T C V H Q A L K G F T T K G V T S 220

721 GTCTCTCAGA TCTTCCACAG CCCAGACCTG GCCATAAGGG ACACCTTTGT GAATGCCTCT 780
221 V S Q I F H S P D L A I R D T F V N A S 240

781 CGGACCCTGT ACAGCAGCAG CCCAGAGTC CTAAGCAACA ACAGTGACGC CAACTTGGAG 840
241 R T L Y S S S P R V L S N N S D A N L E 260

841 CTCATCAACA CCTGGGTGGC CAAGAACACC AACAACAAGA TCAGCCGGCT GCTAGACAGT 900
261 L I N T W V A K N T N N K I S R L L D S 280

901 CTGCCCTCCG ATACCCGCCT TGTCTCCTC AATGCTATCT ACCTGAGTGC CAAGTGAAG 960
281 L P S D T R L V L L N A I Y L S A K W K 300

```

Figure 13B

```

961 ACAACATTG· ATCCCAAGAA· AACCAGAATG· GAACCCTTTC· ACTTCAAAAA· CTCAGTTATA· 1020
301 T T F D P K K T R M E P F H F K N S V I 320

1021 AAAGTGCCCA· TGATGAATAG· CAAGAAGTAC· CCTGTGGCCC· ATTCATTGA· CCAAACCTTG· 1080
321 K V P M M N S K K Y P V A H F I D Q T L 340

1081 AAAGCCAAGG· TGGGGCAGCT· GCAGCTCTCC· CACAATCTGA· GTTGGGTGAT· CCTGGTACCC· 1140
341 K A K V G Q L Q L S H N L S L V I L V P 360

1141 CAGAACCTGA· AACATCGTCT· TGAAGACATG· GAACAGGCTC· TCAGCCCTTC· TGTTTTCAAG· 1200
361 Q N L K H R L E D M E Q A L S P S V F K 380

1201 GCCATCATGG· AGAAACTGGA· GATGTCCAAG· TTCCAGCCCA· CTCTCCTAAC· ACTACCCCGC· 1260
381 A I M E K L E M S K F Q P T L L T L P R 400

1261 ATCAAAGTGA· CGACCAGCCA· GGATATGCTC· TCAATCATGG· AGAAATTGGA· ATTCTTCGAT· 1320
401 I K V T T S Q D M L S I M E K L E F F D 420

1321 TTTCTTATG· ACCTTAACCT· GTGTGGGCTG· ACAGAGGACC· CAGATCTTCA· GGTTCCTGCG· 1380
421 F S Y D L N L C G L T E D P D L Q V S A 440

1381 ATGCAGCACC· AGACAGTGCT· GGAAC TGACA· GAGACTGGGG· TGGAGGCGGC· TGCAGCCTCC· 1440
441 M Q H Q T V L E L T E T G V E A A A A S 460

1441 GCCATCTCTG· TGGCCCGCAC· CCTGCTGGTC· TTTGAAGTGC· AGCAGCCCTT· CCTCTTCGTG· 1500
461 A I S V A R T L L V F E V Q Q P F L F V 480

1501 CTCTGGGACC· AGCAGCACAA· GTTCCCTGTC· TTCATGGGGC· GAGTATATGA· CCCCAGGGCC· 1560
481 L W D Q Q H K F P V F M G R V Y D P R A 500

1561 TGAGACCTGC· AGGATCAGGT· TAGGGCGAGC· GCTACCTCTC· CAGCCTCAGC· TCTCAGTTGC· 1620

1621 AGCCCTGCTG· CTGCCTGCCT· GGA CTTGCCC· CTGCCACCTC· CTGCCTCAGG· TGTCCGCTAT· 1680

1681 CCACCAAAAG· GGCTCCTGAG· GGTCTGGGCA· AGGGACCTGC· TTCTATTAGC· CCTTCTCCAT· 1740

1741 GGCCCTGCCA· TGCTCTCCAA· ACCACTTTTT· GCAGCTTTCT· CTAGTTCAAG· TTCACCAGAC· 1800

1801 TCTATAAATA· AAACCTGACA· GACCAT 1826

```


Figure 14A

```

1 AGTCTGCACT GGAGCTGCCT GGTGACCAGA AGTTTGAGT CCGCTGACGT CGCCGCCCAG 60

61 ATGGCCTCCA GGCTGACCCT GCTGACCCTC CTGCTGCTGC TGCTGGCTGG GGATAGAGCC 120
1 M A S R L T L L T L L L L L A G D R A 20

121 TCCTCAAATC CAAATGCTAC CAGCTCCAGC TCCCAGGATC CAGAGAGTTT GCAAGACAGA 180
21 S S N P N A T S S S S Q D P E S L Q D R 40

181 GGCGAAGGGA AGGTCGCAAC AACAGTTATC TCCAAGATGC TATTCGTTGA ACCCATCCTG 240
41 G E G K V A T T V I S K M L F V E P I L 60

241 GAGGTTTCCA GCTTGCCGAC AACCAACTCA ACAACCAATT CAGCCACCAA AATAACAGCT 300
61 E V S S L P T T N S T T N S A T K I T A 80

301 AATACCACTG ATGAACCCAC CACACAACCC ACCACAGAGC CCACCACCCA ACCCACCATC 360
81 N T T D E P T T Q P T T E P T T Q P T I 100

361 CAACCCACCC AACCAACTAC CCAGCTCCCA ACAGATTCTC CTACCCAGCC CACTACTGGG 420
101 Q P T Q P T T Q L P T D S P T Q P T T G 120

421 TCCTTCTGCC CAGGACCTGT TACTCTCTGC TCTGACTTGG AGAGTCATTC AACAGAGGCC 480
121 S F C P G P V T L C S D L E S H S T E A 140

481 GTGTTGGGGG ATGCTTTGGT AGATTTCTCC CTGAAGCTCT ACCACGCCTT CTCAGGAATG 540
141 V L G D A L V D F S L K L Y H A F S G M 160

541 AAGAAGGTGG AGACCAACAT GGCCTTTTCC CCATTCAACA TCGCCAGCCT CCTTACCCAG 600
161 K K V E T N M A F S P F S I A S L L T Q 180

601 GTCTGCTCG GGGCTGGGCA GAACACCAAA ACAAACCTGG AGAGCATCCT CTCTTACCCC 660
181 V L L G A G Q N T K T N L E S I L S Y P 200

661 AAGGACTTCA CCTGTGTCCA CCAGGCCCTG AAGGGCTTCA CGACCAAAGG TGTCACCTCA 720
201 K D F T C V H Q A L K G F T T K G V T S 220

721 GTCTCTCAGA TCTTCCACAG CCCAGACCTG GCCATAAGGG ACACCTTTGT GAATGCCTCT 780
221 V S Q I F H S P D L A I R D T F V N A S 240

781 CGGACCCTGT ACAGCAGCAG CCCAGAGTC CTAAGCAACA ACAGTGACGC CAACTTGGAG 840
241 R T L Y S S S P R V L S N N S D A N L E 260

841 CTCATCAACA CCTGGGTGGC CAAGAACACC AACAACAAGA TCAGCCGGCT GCTAGACAGT 900
261 L I N T W V A K N T N N K I S R L L D S 280

901 CTGCCCTCCG ATACCCGCCT TGTCTCCTC AATGCTATCT ACCTGAGTGC CAAGTGAAG 960
281 L P S D T R L V L L N A I Y L S A K W K 300

```

Figure 14B

```

961 ACAACATTG ATCCCAAGAA AACCAGAATG GAACCCTTTC ACTTCAAAAA CTCAGTTATA 1020
301 T T F D P K K T R M E P F H F K N S V I 320

1021 AAAGTGCCCA TGATGAATAG CAAGAAGTAC CCTGTGGCCC ATTCATTGA CCAAACCTTG 1080
321 K V P M M N S K K Y P V A H F I D Q T L 340

1081 AAAGCCAAGG TGGGGCAGCT GCAGCTCTCC CACAATCTGA GTTTGGTGAT CCTGGTACCC 1140
341 K A K V G Q L Q L S H N L S L V I L V P 360

1141 CAGAACCTGA AACATCGTCT TGAAGACATG GAACAGGCTC TCAGCCCTTC TGTTCCTCAAG 1200
361 Q N L K H R L E D M E Q A L S P S V F K 380

1201 GCCATCATGG AGAAACTGGA GATGTCCAAG TTCCAGCCCA CTCTCCTAAC ACTACCCCGC 1260
381 A I M E K L E M S K F Q P T L L T L P R 400

1261 ATCAAAGTGA CGACCAGCCA GGATATGCTC TCAATCATGG AGAAATTGGA ATTCTTCGAT 1320
401 I K V T T S Q D M L S I M E K L E F F D 420

1321 TTTTCTTATG ACCTTAACCT GTGTGGGCTG ACAGAGGACC CAGATCTTCA GGTTTCTGCG 1380
421 F S Y D L N L C G L T E D P D L Q V S A 440

1381 ATGCAGCACC AGACAGTGCT GGAAGTACA GAGACTGGGG TGGAGGCGGC TGCAGCCTCC 1440
441 M Q H Q T V L E L T E T G V E A A A A S 460

1441 GCCATCTCTG TGGCCCGCAC CCTGCTGGTC TTTGAAGTGC AGCAGCCCTT CCTCTTCGTG 1500
461 A I S V A R T L L V F E V Q Q P F L F V 480

1501 CTCTGGGACC AGCAGCACAA GTTCCCTGTC TTCATGGGGC GAGTATATGA CCCCAGGGCC 1560
481 L W D Q Q H K F P V F M G R V Y D P R A 500

1561 TGAGACCTGC AGGATCAGGT TAGGGCGAGC GCTACCTCTC CAGCCTCAGC TCTCAGTTGC 1620

1621 AGCCCTGCTG CTGCCTGCCT GGAAGTGCCC CTGCCACCTC CTGCCTCAGG TGTCCGCTAT 1680

1681 CCACCAAAAG GGCTCCTGAG GGTCTGGGCA AGGGACCTGC TTCTATTAGC CCTTCTCCAT 1740

1741 GGCCCTGCCA TGCTCTCCAA ACCACTTTTT GCAGCTTTCT CTAGTTCAAG TTCACCAGAC 1800

1801 TCTATAAATA AACCTGACA GACCAT 1826

```

Figure 15A

```

1  AGTCTGCACT  GGAGCTGCCT  GGTGACCAGA  AGTTTGGAGT  CCGCTGACGT  CGCCGCCCCAG  60

61  ATGGCCTCCA  GGCTGACCCT  GCTGACCCTC  CTGCTGCTGC  TGCTGGCTGG  GGATAGAGCC  120
1  M  A  S  R    L  T  L    L  T  L    L  L  L  L    L  A  G    D  R  A    20

121 TCCTCAAATC  CAAATGCTAC  CAGCTCCAGC  TCCCAGGATC  CAGAGAGTTT  GCAAGACAGA  180
21 S  S  N  P    N  A  T    S  S  S    S  Q  D  P    E  S  L    Q  D  R    40

181 GGCGAAGGGA  AGGTCGCAAC  AACAGTTATC  TCCAAGATGC  TATTCGTTGA  ACCCATCCTG  240
41 G  E  G  K    V  A  T    T  V  I    S  K  M  L    F  V  E    P  I  L    60

241 GAGGTTTCCA  GCTTGCCGAC  AACCAACTCA  ACAACCAATT  CAGCCACCAA  AATAACAGCT  300
61 E  V  S  S    L  P  T    T  N  S    T  T  N  S    A  T  K    I  T  A    80

301 AATACCACTG  ATGAACCCAC  CACACAACCC  ACCACAGAGC  CCACCACCCA  ACCCACCATC  360
81 N  T  T  D    E  P  T    T  Q  P    T  T  E  P    T  T  Q    P  T  I    100

361 CAACCCACCC  AACCAACTAC  CCAGCTCCCA  ACAGATTCTC  CTACCCAGCC  CACTACTGGG  420
101 Q  P  T  Q    P  T  T    Q  L  P    T  D  S  P    T  Q  P    T  T  G    120

421 TCCTTCTGCC  CAGGACCTGT  TACTCTCTGC  TCTGACTTGG  AGAGTCATTC  AACAGAGGCC  480
121 S  F  C  P    G  P  V    T  L  C    S  D  L  E    S  H  S    T  E  A    140

481 GTGTTGGGGG  ATGCTTTGGT  AGATTTCTCC  CTGAAGCTCT  ACCACGCCTT  CTCAGCAATG  540
141 V  L  G  D    A  L  V    D  F  S    L  K  L  Y    H  A  F    S  A  M    160

541 AAGAAGGTGG  AGACCAACAT  GGCCTTTTCC  CCATTTCAGC  TCGCCAGCCT  CCTTACCCAG  600
161 K  K  V  E    T  N  M    A  F  S    P  F  S  I    A  S  L    L  T  Q    180

601 GTCCTGCTCG  GGGCTGGGCA  GAACACCAAA  ACAACCTGG  AGAGCATCCT  CTCTTACCCC  660
181 V  L  L  G    A  G  Q    N  T  K    T  N  L  E    S  I  L    S  Y  P    200

661 AAGGACTTCA  CCTGTGTCCA  CCAGGCCCTG  AAGGGCTTCA  CGACCAAAGG  TGTACCTCA  720
201 K  D  F  T    C  V  H    Q  A  L    K  G  F  T    T  K  G    V  T  S    220

721 GTCTCTCAGA  TCTTCCACAG  CCCAGACCTG  GCCATAAGGG  ACACCTTTGT  GAATGCCTCT  780
221 V  S  Q  I    F  H  S    P  D  L    A  I  R  D    T  F  V    N  A  S    240

781 CGGACCCTGT  ACAGCAGCAG  CCCCAGAGTC  CTAAGCAACA  ACAGTGACGC  CAACTTGGAG  840
241 R  T  L  Y    S  S  S    P  R  V    L  S  N  N    S  D  A    N  L  E    260

841 CTCATCAACA  CCTGGGTGGC  CAAGAACACC  AACAACAAGA  TCAGCCGGCT  GCTAGACAGT  900
261 L  I  N  T    W  V  A    K  N  T    N  N  K  I    S  R  L    L  D  S    280

901 CTGCCCTCCG  ATACCCGCCT  TGTCTCTCTC  AATGCTATCT  ACCTGAGTGC  CAAGTGAAG  960
281 L  P  S  D    T  R  L    V  L  L    N  A  I  Y    L  S  A    K  W  K    300

```

Figure 15B

```

961 ACAACATTTG ATCCCAAGAA AACCAGAATG GAACCCTTTC ACTTCAAAAA CTCAGTTATA 1020
301 T T F D P K K T R M E P F H F K N S V I 320

1021 AAAGTGCCCA TGATGAATAG CAAGAAGTAC CCTGTGGCCC ATTTCATTGA CCAAACCTTG 1080
321 K V P M M N S K K Y P V A H F I D Q T L 340

1081 AAAGCCAAGG TGGGGCAGCT GCAGCTCTCC CACAATCTGA GTTGGGTGAT CCTGGTACCC 1140
341 K A K V G Q L Q L S H N L S L V I L V P 360

1141 CAGAACCTGA AACATCGTCT TGAAGACATG GAACAGGCTC TCAGCCCTTC TGTTTTCAAG 1200
361 Q N L K H R L E D M E Q A L S P S V F K 380

1201 GCCATCATGG AGAAACTGGA GATGTCCAAG TTCCAGCCCA CTCTCCTAAC ACTACCCCGC 1260
381 A I M E K L E M S K F Q P T L L T L P R 400

1261 ATCAAAGTGA CGACCAGCCA GGATATGCTC TCAATCATGG AGAAATTGGA ATTCTTCGAT 1320
401 I K V T T S Q D M L S I M E K L E F F D 420

1321 TTTTCTTATG ACCTTAACCT GTGTGGGCTG ACAGAGGACC CAGATCTTCA GGTTCCTGCG 1380
421 F S Y D L N L C G L T E D P D L Q V S A 440

1381 ATGCAGCACC AGACAGTGCT GGAAGTGACA GAGACTGGGG TGGAGGCGGC TGCAGCCTCC 1440
441 M Q H Q T V L E L T E T G V E A A A A S 460

1441 GCCATCTCTG TGGCCCGCAC CCTGCTGGTC TTTGAAGTGC AGCAGCCCTT CCTCTTCATG 1500
461 A I S V A R T L L V F E V Q Q P F L F M 480

1501 CTCTGGGACC AGCAGCACAA GTTCCCTGTC TTCATGGGGC GAGTATATGA CCCAGGGGCC 1560
481 L W D Q Q H K F P V F M G R V Y D P R A 500

1561 TGAGACCTGC AGGATCAGGT TAGGGCGAGC GCTACCTCTC CAGCCTCAGC TCTCAGTTGC 1620

1621 AGCCCTGCTG CTGCCTGCCT GGAAGTGCCC CTGCCACCTC CTGCCTCAGG TGTCCGCTAT 1680

1681 CCACCAAAAG GGCTCCTGAG GGTCTGGGCA AGGGACCTGC TTCTATTAGC CCTTCTCCAT 1740

1741 GGCCCTGCCA TGCTCTCCAA ACCACTTTTT GCAGCTTTCT CTAGTTCAAG TTCACCAGAC 1800

1801 TCTATAAATA AAACCTGACA GACCAT 1826

```

Figure 16

```

1 TCCTCCACCT GCTGGCCCTT GGACACCTCT GTCACCATGT GGTTCCTGGT TCTGTGCCTC 60
1 M W F L V L C L 8

61 GCCCTGTCCC TGGGGGGGAC TGGTGCTGCG CCCCCGATTC AGTCCCGGAT TGTGGGAGGC 120
9 A L S L G G T G A A P P I Q S R I V G G 28

121 TGGGAGTGTG AGCAGCATTC CCAGCCCTGG CAGGCGGCTC TGTACCATT T CAGCACTTTC 180
29 W E C E Q H S Q P W Q A A L Y H F S T F 48

181 CAGTGTGGGG GCATCCTGGT GCACCGCCAG TGGTGCTCA CAGCTGCTCA TTGCATCAGC 240
49 Q C G G I L V H R Q W V L T A A H C I S 68

241 GACAATTACC AGCTCTGGCT GGGTCGCCAC AACTTGTTTG ACGACGAAAA CACAGCCCAG 300
69 D N Y Q L W L G R H N L F D D E N T A Q 88

301 TTGTTCATG TCAGTGAGAG CTTCCACAC CTTGGCTTCA ACATGAGCCT CCTGGAGAAC 360
89 F V H V S E S F P H P G F N M S L L E N 108

361 CACACCCGCC AAGCAGACGA GGACTACAGC CACGACCTCA TGCTGCTCCG CCTGACAGAG 420
109 H T R Q A D E D Y S H D L M L L R L T E 128

421 CCTGCTGATA CCATCACAGA TGCTGTGAAG GTCGTGGAGT TGCCACCGA GGAACCCGAA 480
129 P A D T I T D A V K V V E L P T E E P E 148

481 GTGGGGAGCA CCTGTTTGGC TTCCGGCTGG GGCAGCATCG AACCAGAGAA TTTCTCATTT 540
149 V G S T C L A S G W G S I E P E N F S F 168

541 CCAGATGATC TCCAGTGTGT GGACCTCAA ATCCTGCCTA ATGATGAGTG CAAAAAGCC 600
169 P D D L Q C V D L K I L P N D E C K K A 188

601 CACGTCCAGA AGGTGACAGA CTTCATGCTG TGTGTCGGAC ACCTGGAAGG TGGCAAAGAC 660
189 H V Q K V T D F M L C V G H L E G G K D 208

661 ACCTGTGTGG GTGATTCAGG GGGCCCGCTG ATGTGTGATG GTGTGCTCCA AGGTGTCACA 720
209 T C V G D S G G P L M C D G V L Q G V T 228

721 TCATGGGGCT ACGTCCCTTG TGGCACCCC AATAAGCCTT CTGTCGCCGT CAGAGTGCTG 780
229 S W G Y V P C G T P N K P S V A V R V L 248

781 TCTTATGTGA AGTGGATCGA GGACACCATA GCGGAGAACT CCTGAACGCC CAGCCCTGTC 840
249 S Y V K W I E D T I A E N S * 262

841 CCCTACCCCC AGTAAAATCA AATGTGCATC C 871

```

Figure 17

```

1 TCCTCCACCT GCTGGCCCCCT GGACACCTCT GTCACCATGT GGTTCCTGGT TCTGTGCCTC 60
1 M W F L V L C L 8

61 GCCCTGTCCC TGGGGGGGAC TGGTGTGTCG CCCCCGATTC AGTCCCGGAT TGTGGGAGGC 120
9 A L S L G G T G A A P P I Q S R I V G G 28

121 TGGGAGTGTG AGCAGCATTC CCAGCCCTGG CAGGCGGCTC TGTACCATT T CAGCACTTTC 180
29 W E C E Q H S Q P W Q A A L Y H F S T F 48

181 CAGTGTGGGG GCATCCTGGT GCACCGCCAG TGGGTGCTCA CAGCTGCTCA TTGCATCAGC 240
49 Q C G G I L V H R Q W V L T A A H C I S 68

241 GACAATTACC AGCTCTGGCT GGGTCGCCAC AACTTGTTTG ACGACGAAAA CACAGCCCAG 300
69 D N Y Q L W L G R H N L F D D E N T A Q 88

301 TTGTTTCATG TCAGTGAGAG CTTCCCACAC CCTGGCTTCA ACATGAGCCT CCTGGAGAAC 360
89 F V H V S E S F P H P G F N M S L L E N 108

361 CACACCCGCC AAGCAGACGA GGACTACAGC CACGACCTCA TGCTGCTCCG CCTGACAGAG 420
109 H T R Q A D E D Y S H D L M L L R L T E 128

421 CCTGCTGATA CCATCACAGA TGCTGTGAAG GTCGTGGAGT TGCCCAACCGA GGAACCCGAA 480
129 P A D T I T D A V K V V E L P T E E P E 148

481 GTGGGGAGCA CCTGTTTGGC TTCCGGCTGG GGCAGCATCG AACCAGAGAA TTTCTCATTT 540
149 V G S T C L A S G W G S I E P E N F S F 168

541 CCAGATGATC TCCAGTGTGT GGACCTCAA A TCCTGCCTA ATGATGAGTG CGAAAAAGCC 600
169 P D D L Q C V D L K I L P N D E C E K A 188

601 CACGTCCAGA AGGTGACAGA CTTCATGCTG TGTGTGCGAC ACCTGGAAGG TGGCAAAGAC 660
189 H V Q K V T D F M L C V G H L E G G K D 208

661 ACCTGTGTGG GTGATTCAGG GGGCCCCTG ATGTGTGATG GTGTGCTCCA AGGTGTCACA 720
209 T C V G D S G G P L M C D G V L Q G V T 228

721 TCATGGGGCT ACGTCCCTTG TGGCACCCCT AATAAGCCTT CTGTCGCCGT CAGAGTGCTG 780
229 S W G Y V P C G T P N K P S V A V R V L 248

781 TCTTATGTGA AGTGGATCGA GGACACCATA GCGGAGAACT CCTGAACGCC CAGCCCTGTC 840
249 S Y V K W I E D T I A E N S * 262

841 CCCTACCCCC AGTAAAATCA AATGTGCATC C 871

```

Figure 18

```

1 TCCTCCACCT GCTGGCCCCCT GGACACCTCT GTCACCATGT GGTTCCTGGT TCTGTGCCTC 60
1 M W F L V L C L 8

61 GCCCTGTCCC TGGGGGGGAC TGGTGCTGCG CCCCCGATTC AGTCCCGGAT TGTGGGAGGC 120
9 A L S L G G T G A A P P I Q S R I V G G 28

121 TGGGAGTGTG AGCAGCATTG CCAGCCCTGG CAGGCGGCTC TGTACCATTG CAGCACTTTC 180
29 W E C E Q H S Q P W Q A A L Y H F S T F 48

181 CAGTGTGGGG GCATCCTGGT GCACCGCCAG TGGGTGCTCA CAGCTGCTCA TTGCATCAGC 240
49 Q C G G I L V H R Q W V L T A A H C I S 68

241 GACAATTACC AGCTCTGGCT GGGTCGCCAC AACTTGTTTG ACGACGAAAA CACAGCCCAG 300
69 D N Y Q L W L G R H N L F D D E N T A Q 88

301 TTTGTTTCATG TCAGTGAGAG CTTCCCACAC CCTGGCTTCA ACATGAGCCT CCTGGAGAAC 360
89 F V H V S E S F P H P G F N M S L L E N 108

361 CACACCCGCC AAGCAGACGA GGACTACAGC CACGACCTCA TGCTGCTCCG CCTGACAGAG 420
109 H T R Q A D E D Y S H D L M L L R L T E 128

421 CCTGCTGATA CCATCACAGA TGCTGTGAAG GTCGTGGAGT TGCCACCCCA GGAACCCGAA 480
129 P A D T I T D A V K V V E L P T Q E P E 148

481 GTGGGGAGCA CCTGTTTGGC TTCCGGCTGG GGCAGCATCG AACCAGAGAA TTTCTCATTT 540
149 V G S T C L A S G W G S I E P E N F S F 168

541 CCAGATGATC TCCAGTGTGT GGACCTCAA AATCCTGCCTA ATGATGAGTG CAAAAAGCC 600
169 P D D L Q C V D L K I L P N D E C K K A 188

601 CACGTCCAGA AGGTGACAGA CTTGATGCTG TGTGTCGGAC ACCTGGAAGG TGGCAAAGAC 660
189 H V Q K V T D F M L C V G H L E G G K D 208

661 ACCTGTGTGG GTGATTCAGG GGGCCCGCTG ATGTGTGATG GTGTGCTCCA AGGTGTCACA 720
209 T C V G D S G G P L M C D G V L Q G V T 228

721 TCATGGGGCT ACGTCCCTTG TGGCACCCC AATAAGCCTT CTGTCGCCGT CAGAGTGCTG 780
229 S W G Y V P C G T P N K P S V A V R V L 248

781 TCTTATGTGA AGTGGATCGA GGACACCATA GCGGAGAACT CCTGAACGCC CAGCCCTGTC 840
249 S Y V K W I E D T I A E N S * 262

841 CCCTACCCCC AGTAAAATCA AATGTGCATC C 871

```

Figure 19

BRB1_MOUSE MASQ.ASLKLPNSQSQAPPNITSCEGAPEAWDLLCRVLPGFVITICFFGLLGNLLVLS
 BRB1_RAT MASE.VLLELPNSRSLQAPANITSCESALEDWLLYRVLPGFVITICFFGLLGNLLVLS
 BRB1_HUMAN MASSWPPELQSSNQSLFPQATACDPAEAWDLLHRVLPFTTISICFFGLLGNLFVLL
 BRB1_RABIT MASQ.GPLELPNSQSQLAPPNATSCSGAPDAWDLHRLPFTTIAIFTLLGNGNSFVLS

BRB1_MOUSE FFLLPWRRWW...QORRQRLTIAEIYLANLAASDLVFVLGLPFWAENVGNRFNWPFGS DL
 BRB1_RAT FFLLPWRQWWQQRQORRQRLTIAEIYLANLAASDLVFVLGLPFWAENVGNRFNWPFGTDL
 BRB1_HUMAN VFLLP.....RRQLNVAEIYLANLAASDLVFVLGLPFWAENVGNRFNWPFGALL
 BRB1_RABIT VFLLA.....RRRLSVAEIYLANLAASDLVFVLGLPFWAENVGNRFNWPFGAAL

BRB1_MOUSE CRVVSQVIKANLFISIFLVVAISQDRYRLVYPMTSWGNNRRRRQAQVTCLLIWVAGGLLS
 BRB1_RAT CRVVSQVIKANLFISIFLVVAISQDRYRLVYPMTSWGNYRRRRQAQATCCLLIWVAGGLLS
 BRB1_HUMAN CRVINGVIKANLFISIFLVVAISQDRYRVLVHPMASGRQORRRQAQVTCVLIWVAGGLLS
 BRB1_RABIT CRVINGVIKANLFISIFLVVAISQDRYSVLVHPMASRRGRRRRQAQATCALIWVAGGLLS

BRB1_MOUSE TPTFLRSVKVVPDLNISACILLFPHEAWHFVRMVELNVLGFLPLAAILFFNYHILASL
 BRB1_RAT IPTFLRSVKVVPDLNVISACILLFPHEAWHFARMVELNVLGFLPLVTAIFFNYHILASL
 BRB1_HUMAN IPTFLRSIQAVPDLNITACILLFPHEAWHFARIVELNVLGFLPLAAILFFNYHILASL
 BRB1_RABIT TPTFLRSVRAVPELVNISACILLFPHEAWHFLRMVELNVLGFLPLAAILFFNCHILASL

*

BRB1_MOUSE RCQKEASRTRCGGPKDSKTMGLILTLVASFLVCWAPYHFFAFLEFLVQVRVIQDCFWKEL
 BRB1_RAT RCQKEASRTRCGGPKGSKTTGLILTLVASFLVCWCPYHFFAFLEFLVQVRVIQDCSWKEI
 BRB1_HUMAN RTREEVSRTVRGPKDSKTTALILTLVVAFLVCWAPYHFFAFLEFLVQVQAVRGCFWEDF
 BRB1_RABIT RRRGERVPSRCGGPKDSKSTALILTLVASFLVCWAPYHFFAFLECLVQVHAIGGCFWEEF

*

BRB1_MOUSE TDLGLQLANFFAFVNSCLNPIYVFAGLEKTRVLGTL~~~~~
 BRB1_RAT TDLGLQLANFFAFVNSCLNPIYVFAGLEKTRVLGTL~~~~~
 BRB1_HUMAN IDLGLQLANFFAFVNSCLNPIYVFVGLFKTRVWELYKQCTPKSLAPSSSHRKELFQL
 BRB1_RABIT TDLGLQLSNFSAFVNSCLNPIYVFVGLFKTRVWELCQCCSPKSLAPSSSRKEMLWG

BRB1_MOUSE ~~~~~
 BRB1_RAT ~~~~~
 BRB1_HUMAN FWRN
 BRB1_RABIT FWRN

Figure 20

BRB2_MOUSE ~~~~~MPCSWKLLGFLSVHE.PMPTAASFGIEMFNVTTOVLGSALNGTISKDN.CPDTEW
 BRB2_RAT MDTRSSLCP.KTQAVVAVFW.GPGCHLSTCIEMFNITTOALGSAHNGTFSEVN.CPDTEW
 BRB2_RABIT ~~~~~MLNITSQVLAAPALNGSVSOSGCPNTEW
 BRB2_CAVPO ~~~~~MFNITSQV..SALNATLAOGNSCLDAEW
 BRB2_HUMAN ~~~~~MFSPWKISMFLSVREDSVPTTASFADMLNVTLO..GPTLNGTFAQ.SKCPQVEW

BRB2_MOUSE WSWLNATQAPFLWVLFVLAALENIFVLSVFVFLHKNSCTVAEIYLGNLAAADLILACGLPF
 BRB2_RAT WSWLNATQAPFLWVLFVLAALENIFVLSVFCLHKINCTVAEIYLGNLAAADLILACGLPF
 BRB2_RABIT SGWLNVIQAPFLWVLFVLALENIFVLSVFCLHKSSCTVAEIYLGNLAAADLILACGLPF
 BRB2_CAVPO WSWLNATQAPFLWVLFVLALENIFVLSVFVFLHKSSCTVAEIYLGNLAAADLILAFGLPF
 BRB2_HUMAN LGWLNATQAPFLWVLFVLALENIFVLSVFCLHKSSCTVAEIYLGNLAAADLILACGLPF

BRB2_MOUSE WAITIANNFDWVFGVLCRVVNTMIYMNLYSSICFLMLVSIIDRYLALVKTMSMGRMRGVR
 BRB2_RAT WAITIANNFDWVFGVLCRVVNTMIYMNLYSSICFLMLVSIIDRYLALVKTMSMGRMRGVR
 BRB2_RABIT WAITIANHFDWVFGVLCRVVNTMIYMNLYSSICFLMLVSIIDRYLALVKTMSGRMRVR
 BRB2_CAVPO WAITIANNFDWVFGVLCRVVNTMIYMNLYSSICFLMLVSIIDRYLALVKTMSMGRMRGVR
 BRB2_HUMAN WAITISNNFDWVFGVLCRVVNTMIYMNLYSSICFLMLVSIIDRYLALVKTMSMGRMRGVR

BRB2_MOUSE WAKLYSLVIWGCTLLSSPMLVFRTMKREYSEEGHNVTACVIVYPSRSWEVFTNVLLNLVG
 BRB2_RAT WAKLYSLVIWCTLLSSPMLVFRTMKDYREEGHNVTACVIVYPSRSWEVFTNVLLNLVG
 BRB2_RABIT WAKLYSLVIWGCTLLSSPMLVFRTMKDYRDEGYNVTACIDYPSRSWEVFTNVLLNLVG
 BRB2_CAVPO WAKLYSLVIWGCTLLSSPMLVFRTMKDYRDEGHNVTACIDYPSLIWQVFTNVLLNLVG
 BRB2_HUMAN WAKLYSLVIWGCTLLSSPMLVFRTMKREYSEEGHNVTACVISYPSLIWEVFTNVLLNLVG

BRB2_MOUSE FLLPLSVITFCTVRIQVLRNEMKKFKFQTERKATVVLAVLGLFVLCWVPFQISTFL
 BRB2_RAT FLLPLSVITFCTVRIQVLRNEMKKFKFQTERKATVVLAVLGLFVLCWVPFQISTFL
 BRB2_RABIT FLLPLSVITFCTVQIQVLRNEMQKFKEIQTERRATVVLAVLLLFVVCWLPFQISTFL
 BRB2_CAVPO FLLPLSVITFCTVQIQVLRNEMQKFKEIQTERRATVVLAVLLLFVVCWLPFQIGTFL
 BRB2_HUMAN FLLPLSVITFCTVQIQVLRNEMQKFKEIQTERRATVVLAVLLLFVVCWLPFQISTFL

BRB2_MOUSE DTLLRLGVLSCGWDEHVIDVITQISSYVAYSNSCLNPLVYVIVGKFRKKSREVYRVLCQ
 BRB2_RAT DTLLRLGVLSCGWNERAVDVIDVITQISSYVAYSNSCLNPLVYVIVGKFRKKSREVYQALCR
 BRB2_RABIT DTLLRLGVLSCGWDEHVIDVITQISSYVAYSNSCLNPLVYVIVGKFRKKSREVYRAACP
 BRB2_CAVPO DTLRLGVLSCGWDEHVIDVITQISSYVAYSNSCLNPLVYVIVGKFRKKSREVYHGLCR
 BRB2_HUMAN DTLRLGVLSCGWDEHVIDVITQIASVAYSNSCLNPLVYVIVGKFRKKSWEVYQGVCO

*

BRB2_MOUSE KGGCMGEPVQMENSMTGLRTSISVERQIHKLQDWAGKKQ~~~~~
 BRB2_RAT KGGCMGESVQMENSMTGLRTSISVERQIHKLQDWAGNKKQ~~~~~
 BRB2_RABIT KAGCVLEPVQAESSMTGLRTSISVERQIHKLPENWTRSSQ~~~~~
 BRB2_CAVPO SGGCVSEPAQSENSMTGLRTSISVERQIHKLQDWARSSSEGTTPPGLL
 BRB2_HUMAN KGGCRSEPIQMENSMTGLRTSISVERQIHKLQDWAGSRQ~~~~~

Figure 21A

```

1 CTGTGCATGG CATCATCCTG GCCCCCTCTA GAGCTCCAAT CCTCCAACCA GAGCCAGCTC 60
1      M A S S W P P L E L Q S S N Q S Q L 18

61 TTCCCTCAAA ATGCTACGGC CTGTGACAAT GCTCCAGAAG CCTGGGACCT GCTGCACAGA 120
19 F P Q N A T A C D N A P E A W D L L H R 38

121 GTGCTGCCNA CATTATCAT CTCCATCTGT TTCTTCGGCC TCCTAGGGAA CCTTTTTGTC 180
39 V L P T F I I S I C F F G L L G N L F V 58

181 CTGTTGGTCT TCCTCCTGCC CCGGCGGCAA CTGAACGTGG CAGAAATCTA CCTGGCCAAC 240
59 L L V F L L P R R Q L N V A E I Y L A N 78

241 CTGGCAGCCT CTGATCTGGT GTTTGTCTTG GGCTTGCCCT TCTGGGCAGA GAATATCTGG 300
79 L A A S D L V F V L G L P F W A E N I W 98

301 AACCAGTTTA ACTGGCCTTT CGGAGCCCTC CTCTGCCGTG TCATCAANGG GGTCATCAAG 360
99 N Q F N W P F G A L L C R V I N G V I K 118

361 GCCAATTTGT TCATCAGCAT CTTCTTGGTG GTGGCCATCA GCCAGGACCG CTACCGCGTG 420
119 A N L F I S I F L V V A I S Q D R Y R V 138

421 CTGGTGCACC CTATGGCCAG CGGAAGGCAG CAGCGGCGGA GNCAGGCCCC GGTCACCTGC 480
139 L V H P M A S G R Q Q R R R Q A R V T C 158

481 GTGCTCATCT GGGTTGTGGG GGGCCTCTTG AGCATCCCCA CATTCCTGCT GCGATCCATC 540
159 V L I W V V G G L L S I P T F L L R S I 178

541 CAAGCCGTCC CAGATCTGAA CATCACCGCC TGCATCNTGC TCCTCCCCCA TGAGGCCTGG 600
179 Q A V P D L N I T A C I X L L P H E A W 198

601 CACTTTGCAA GGATTGTGGA GTTAAATATT CTGGGTTTCC TCCTACCACT GGCTGCGATC 660
199 H F A R I V E L N I L G F L L P L A A I 218

661 GTCTTCTTCA ACTACCACAT CCTGGCCTCC CTGCGAACGC GGGANGAGGT CAGCAGGACA 720
219 V F F N Y H I L A S L R T R E E V S R T 238

721 AGAGTGCNNGG GGCCGAAGGA TAGCAAGACC ACAGCGCTGA TCCTCACGCT CGTGGTTGCC 780
239 R V X G P K D S K T T A L I L T L V V A 258

781 TTCCTGGTCT GCTGGGCCCC TTACCACTTC TTTGCCTTCC TGGAATTCTT ATTCCAGGTG 840
259 F L V C W A P Y H F F A F L E F L F Q V 278

841 CAAGCAGTCC GAGGCTGCTT TTGGGAGGAC TTCATTGACC TGGGCCTGCA ATTGGCCAAC 900
279 Q A V R G C F W E D F I D L G L Q L A N 298

```

Figure 21B

901 TTCTTTGCCT TCACTAACAG CTCCTGAAT CCAGTAATTT ATGTCTTTGT GGGCCNGCTC 960
299 F F A F T N S S L N P V I Y V F V G **X** L 318

961 TTCAGGACCA AGGTCTGGGA ACTTTATAAA CAATGCACCC CTAAAAGTCT TGCTCCAATA 1020
319 F R T K V W E L Y K Q C T P K S L A P I 338

1021 TCTTCATCCC ATAGGAAAGA AATCTTCCAA CTTTCTGGC GGAATTAAAA CAGCATTGAA 1080
339 S S S H R K E I F Q L F W R N * 353

1081 CC 1082

Figure 22A

1 AATTCAGAGC CACCGCGGGC AGGCGGGCAG TGCATCCAGA AGCGTTTATA TTCTGAGCGC 60
 61 CAGTTCAGCT TTCAAAAAGA GTGCTGCCCA TAAAAAGCCT TCCACCCTCC TGTCTGCTTT 120
 121 AGAAGGACCC TGAGCCCCAG GCGCCAGCCA CAGGACTCTG CTGCAGAGGG GGGTTGTGTA 180
 181 CAGATAGTAG GCTTTACGCC TAGCTTCGAA ATGGATAACG TCCTCCCGGT GGA CTGAGAC 240
 1 M D N V L P V D S D 10
 241 CTCTCCCCAA ACATCTCCAC TAACACCTCG GAACCCAATC AGTTCGTGCA ACCAGCCTGG 300
 11 L S P N I S T N T S E P N Q F V Q P A W 30
 301 CAAATTGTCC TTTGGGCAGC TGCCTACACG GTCATTGTGG TGACCTCTGT GGTGGGCAAC 360
 31 Q I V L W A A A Y T V I V V T S V V G N 50
 361 GTGGTAGTGA TGTGGATCAT CTTAGCCCAC AAAAGAATGA GGACAGTGAC GAACTATTTT 420
 51 V V V M W I I L A H K R M R T V T N Y F 70
 421 CTGGTGAACC TGGCCTTCGC GGAGGCCTCC ATGGCTGCAT TCAATACAGT GGTGAACCTC 480
 71 L V N L A F A E A S M A A F N T V V N F 90
 481 ACCTATGCTG TCCACAACGA ATGGTACTAC GGCCTGTCTT ACTGCAAGTT CCACAACCTC 540
 91 T Y A V H N E W Y Y G L F Y C K F H N F 110
 541 TTNCCCATCG CCGCTGTCTT CGCCAGTATC TACTCCATGA CGGCTGTGGC CTTTGATAGG 600
 111 F P I A A V F A S I Y S M T A V A F D R 130
 601 TACATGGCCA TCATACATCC CCTCCAGCCC CGGCTGTCAG CCACAGCCAC CAAAGTGGTC 660
 131 Y M A I I H P L Q P R L S A T A T K V V 150
 661 ATCTGTGTCA TTTGGGTCCT GGCTCTCCTG CTGGCCTTCC CCCAGGGCTA CTACTCAACC 720
 151 I C V I W V L A L L L A F P Q G Y Y S T 170
 721 ACAGAGACCA TGCCAGCAG AGTCGTGTGC ATGATCGAAT GGCCAGAGCA TCCGAACAAG 780
 171 T E T M P S R V V C M I E W P E H P N K 190
 781 ATTTATGAGA AAGTGATACCA CATCTGTGTG ACTGTGCTGA TCTACTTCCT CCCCTGCTG 840
 191 I Y E K V Y H I C V T V L I Y F L P L L 210
 841 GTGATTGGCT ATGCATACAC CGTAGTGGGA ATCACACTAT GGGCCAGTGA GATCCCCGGG 900
 211 V I G Y A Y T V V G I T L W A S E I P G 230
 901 GACTCCTCTG ACCGCTACCA CGAGCAAGTC TCTGCCAAGC GCAAGGTGGT CAAAATGATG 960
 231 D S S D R Y H E Q V S A K R K V V K M M 250

Figure 22B

```

961 ATTGTCGTGG TGTGCACCTT CGCCATCTGC TGGCTGCCCT TCCACATCTT CTCCTCCTG 1020
251 I V V V C T F A I C W L P F H I F F L L 270

1021 CCCTACATCA ACCCAGATCT CTACCTGAAG AAGTTTATCC AGCAGGTCTA CCTGGCCATC 1080
271 P Y I N P D L Y L K K F I Q Q V Y L A I 290

1081 ATGTGGCTGG CCATGAGCTC CACCATGTAC AACCCCATCA TCTACTGCTG CCTCAATGAC 1140
291 M W L A M S S T M Y N P I I Y C C L N D 310

1141 AGGTTCGTC TGGGCTTCAA GCATGCCTTC CGGTGCTGCC CCTTCATCAG CGCCGGCGAC 1200
311 R F R L G F K H A F R C C P F I S A G D 330

1201 TATGAGGGGC TGGAAATGAA ATCCACCCGG TATCTCCAGA CCCAGGGCAG TGTGTACAAA 1260
331 Y E G L E M K S T R Y L Q T Q G S V Y K 350

1261 GTCAGCCGCC TGGAGACCAC CATCTCCACA GTGGTGGGGG CCCACGAGGA GGAGCCAGAG 1320
351 V S R L E T T I S T V V G A H E E E P E 370

1321 GACGGCCCCA AGGCCACACC CTCNTCCCTG GACCTGACCT CCAACTGCTC TTCACGAAGT 1380
371 D G P K A T P S S L D L T S N C S S R S 390

1381 GACTCCAAGA CCATGACAGA GAGCTTCAGC TTCTCCTCCA ATGTGCTCTC CTAGGCCACA 1440
391 D S K T M T E S F S F S S N V L S * 407

1441 GGGCCTTTGG CAGGTGCAGC CCCCCTGCTC TTTGACCTGC CTCCCTTCAT GCATGGAAAT 1500

1501 TCCCTTCATC TGAACCATC AGAAACACCC TCACACTGGG ACTTGCAAAA AGGGTCAGTA 1560

1561 TGGGTAGGG AAAACATTCC ATCCTTGAGT CAAAAAATCT CAATTCTTCC CTATCTTTGC 1620

1621 CACCCTCATG CTGTGTGACT CAAACCAAAT CACTGAACTT TGCTGAGCCT GTAAAAATAA 1680

1681 AGGTCGGACC AGCTTTTCCT CAAGAGCCCA ATGCATTCCA TTTCTGGAAG TGACTTTGGC 1740

1741 TGCATGCGAG TGCTCATTTT AGGATG 1766

```

Figure 23A

```

1  AGTCTGCACT  GGAGCTGCCT  GGTGACCAGA  AGTTTGGAGT  CCGCTGACGT  CGCCGCCCAG  60

61  ATGGCCTCCA  GGCTGACCCT  GCTGACCCTC  CTGCTGCTGC  TGCTGGCTGG  GGATAGAGCC  120
1  M  A  S  R    L  T  L    L  T  L    L  L  L  L    L  A  G    D  R  A    20

121 TCCTCAAATC  CAAATGCTAC  CAGCTCCAGC  TCCCAGGATC  CAGAGAGTTT  GCAAGACAGA  180
21 S  S  N  P    N  A  T    S  S  S    S  Q  D  P    E  S  L    Q  D  R    40

181 GGCGAAGGGA  AGGTCGCAAC  AACAGTTATC  TCCAAGATGC  TATTCGNTGA  ACCCATCCTG  240
41 G  E  G  K    V  A  T    T  V  I    S  K  M  L    F  X  E    P  I  L    60

241 GAGGTTTCCA  GCTTGCCGAC  AACCAACTCA  ACAACCAATT  CAGCCACCAA  AATAACAGCT  300
61 E  V  S  S    L  P  T    T  N  S    T  T  N  S    A  T  K    I  T  A    80

301 AATACCACTG  ATGAACCCAC  CACACAACCC  ACCACAGAGC  CCACCACCCA  ACCCACCATC  360
81 N  T  T  D    E  P  T    T  Q  P    T  T  E  P    T  T  Q    P  T  I    100

361 CAACCCACCC  AACCAACTAC  CCAGCTCCCA  ACAGATTCTC  CTACCCAGCC  CACTACTGGG  420
101 Q  P  T  Q    P  T  T    Q  L  P    T  D  S  P    T  Q  P    T  T  G    120

421 TCCTTCTGCC  CAGGACCTGT  TACTCTCTGC  TCTGACTTGG  AGAGTCATTC  AACAGAGGCC  480
121 S  F  C  P    G  P  V    T  L  C    S  D  L  E    S  H  S    T  E  A    140

481 GTGTTGGGGG  ATGCTTTGGT  AGATTTCTCC  CTGAAGCTCT  ACCACGCCTT  CTCAGNAATG  540
141 V  L  G  D    A  L  V    D  F  S    L  K  L  Y    H  A  F    S  X  M    160

541 AAGAAGGTGG  AGACCAACAT  GGCCTTTTCC  CCATTCAAGC  TCGCCAGCCT  CCTTACCCAG  600
161 K  K  V  E    T  N  M    A  F  S    P  F  S  I    A  S  L    L  T  Q    180

601 GTCTGCTCG  GGGCTGGGCA  GAACACCAA  ACAACCTGG  AGAGCATCCT  CTCTTACCCC  660
181 V  L  L  G    A  G  Q    N  T  K    T  N  L  E    S  I  L    S  Y  P    200

661 AAGGACTTCA  CCTGTGTCCA  CCAGGCCCTG  AAGGGCTTCA  CGACCAAAGG  TGTACCTCA  720
201 K  D  F  T    C  V  H    Q  A  L    K  G  F  T    T  K  G    V  T  S    220

721 GTCTCTCAGA  TCTTCCACAG  CCCAGACCTG  GCCATAAGGG  ACACCTTTGT  GAATGCCTCT  780
221 V  S  Q  I    F  H  S    P  D  L    A  I  R  D    T  F  V    N  A  S    240

781 CGGACCCTGT  ACAGCAGCAG  CCCCAGAGTC  CTAAGCAACA  ACAGTGACGC  CAACTTGGAG  840
241 R  T  L  Y    S  S  S    P  R  V    L  S  N  N    S  D  A    N  L  E    260

841 CTCATCAACA  CCTGGGTGGC  CAAGAACACC  AACAACAAGA  TCAGCCGGCT  GCTAGACAGT  900
261 L  I  N  T    W  V  A    K  N  T    N  N  K  I    S  R  L    L  D  S    280

901 CTGCCCTCCG  ATACCCGCCT  TGTCTCTCTC  AATGCTATCT  ACCTGAGTGC  CAAGTGAAG  960
281 L  P  S  D    T  R  L    V  L  L    N  A  I  Y    L  S  A    K  W  K    300

```

Figure 23B

```

961 ACAACATTG· ATCCAAGAA· AACCAGAATG· GAACCCTTTC· ACTTCAAAAA· CTCAGTTATA· 1020
301 T T F D P K K T R M E P F H F K N S V I 320

1021 AAAGTGCCCA· TGATGAATAG· CAAGAAGTAC· CCTGTGGCCC· ATTCATTGA· CCAAACCTTG· 1080
321 K V P M M N S K K Y P V A H F I D Q T L 340

1081 AAAGCCAAGG· TGGGCGAGCT· GCAGCTCTCC· CACAATCTGA· GTTTGGTGAT· CCTGGTACCC· 1140
341 K A K V G Q L Q L S H N L S L V I L V P 360

1141 CAGAACCTGA· AACATCGTCT· TGAAGACATG· GAACAGGCTC· TCAGCCCTTC· TGTTTTCAAG· 1200
361 Q N L K H R L E D M E Q A L S P S V F K 380

1201 GCCATCATGG· AGAAACTGGA· GATGTCCAAG· TTCCAGCCCA· CTCTCCTAAC· ACTACCCCGC· 1260
381 A I M E K L E M S K F Q P T L L T L P R 400

1261 ATCAAAGTGA· CGACCAGNCA· GGATATGCTC· TCAATCATGG· AGAAATTGGA· ATTCTTCGAT· 1320
401 I K V T T S Q D M L S I M E K L E F F D 420

1321 TTTTCTTATG· ACCTTAACCT· GTGTGGGCTG· ACAGAGGACC· CAGATCTTCA· GGTTCCTGCG· 1380
421 F S Y D L N L C G L T E D P D L Q V S A 440

1381 ATGCAGCACC· AGACAGTGCT· GGAAC TGACA· GAGACTGGGG· TGGAGGCGGC· TGCAGCCTCC· 1440
441 M Q H Q T V L E L T E T G V E A A A A S 460

1441 GCCATCTCTG· TGGCCCGCAC· CCTGCTGGTC· TTTGAAGTGC· AGCAGCCCTT· CCTCTTCNTG· 1500
461 A I S V A R T L L V F E V Q Q P F L F X 480

1501 CTCTGGGACC· AGCAGCACAA· GTTCCCTGTC· TTCATGGGGC· GAGTATATGA· CCCCAGGGCC· 1560
481 L W D Q Q H K F P V F M G R V Y D P R A 500

1561 TGAGACCTGC· AGGATCAGGT· TAGGGCGAGC· GCTACCTCTC· CAGCCTCAGC· TCTCAGTTGC· 1620

1621 AGCCCTGCTG· CTGCCTGCCT· GGACTTGCCC· CTGCCACCTC· CTGCCTCAGG· TGTCGCTAT· 1680

1681 CCACCAAAAG· GGCTCCTGAG· GGTCTGGGCA· AGGGACCTGC· TTCTATTAGC· CCTTCTCCAT· 1740

1741 GGCCCTGCCA· TGCTCTCCAA· ACCACTTTTT· GCAGCTTTCT· CTAGTTCAAG· TTCACCAGAC· 1800

1801 TCTATAAATA· AAACCTGACA· GACCAT 1826

```

Figure 24

```

1 TCCTCCACCT GCTGGCCCCCT GGACACCTCT GTCACCATGT GGTTCCTGGT TCTGTGCCTC 60
1 M W F L V L C L 8

61 GCCCTGTCCC TGGGGGGGAC TGGTGCTGCG CCCCCGATTC AGTCCCGGAT TGTGGGAGGC 120
9 A L S L G G T G A A P P I Q S R I V G G 28

121 TGGGAGTGTG AGCAGCATTC CCAGCCCTGG CAGGCGGCTC TGTACCATT T CAGCACTTTC 180
29 W E C E Q H S Q P W Q A A L Y H F S T F 48

181 CAGTGTGGGG GCATCCTGGT GCACCGCCAG TGGGTGCTCA CAGCTGCTCA TTGCATCAGC 240
49 Q C G G I L V H R Q W V L T A A H C I S 68

241 GACAATTACC AGCTCTGGCT GGGTCGCCAC AACTTGTTTG ACGACGAAAA CACAGCCCAG 300
69 D N Y Q L W L G R H N L F D D E N T A Q 88

301 TTGTTTCATG TCAGTGAGAG CTTCCCACAC CCTGGCTTCA ACATGAGCCT CCTGGAGAAC 360
89 F V H V S E S F P H P G F N M S L L E N 108

361 CACACCCGCC AAGCAGACGA GGACTACAGC CACGACCTCA TGCTGCTCCG CCTGACAGAG 420
109 H T R Q A D E D Y S H D L M L L R L T E 128

421 CCTGCTGATA CCATCACAGA TGCTGTGAAG GTCGTGGAGT TGCCCCACCA GGAACCCGAA 480
129 P A D T I T D A V K V V E L P T X E P E 148

481 GTGGGGAGCA CCTGTTTGGC TTCCGGCTGG GGCAGCATCG AACCAGAGAA TTTCTCATTT 540
149 V G S T C L A S G W G S I E P E N F S F 168

541 CCAGATGATC TCCAGTGTGT GGACCTCAA AATCCTGCCTA ATGATGAGTG CNAAGCC 600
169 P D D L Q C V D L K I L P N D E C X K A 188

601 CACGTCCAGA AGGTGACAGA CTTCATGCTG TGTGTCGGAC ACCTGGAAGG TGGCAAAGAC 660
189 H V Q K V T D F M L C V G H L E G G K D 208

661 ACCTGTGTGG GTGATTCAGG GGGCCCGCTG ATGTGTGATG GTGTGCTCCA AGGTGTCACA 720
209 T C V G D S G G P L M C D G V L Q G V T 228

721 TCATGGGGCT ACGTCCCTTG TGGCACCCCT AATAAGCCTT CTGTCGCCGT CAGAGTGCTG 780
229 S W G Y V P C G T P N K P S V A V R V L 248

781 TCTTATGTGA ACTGGATCGA GGACACCATA GCGGAGAACT CCTGAACGCC CAGCCCTGTC 840
249 S Y V K W I E D T I A E N S * 262

841 CCCTACCCCC AGTAAAATCA AATGTGCATC C 871

```


Figure 25A

```

1 CTGTGCATGG CATCATCCTG GCCCCCTCTA GAGCTCCAAT CCTCCAACCA GAGCCAGCTC 60
1      M A S S W P P L E L Q S S N Q S Q L 18

61 TTCCCTCAAA ATGCTACGGC CTGTGACAAAT GCTCCAGAAG CCTGGGACCT GCTGCACAGA 120
19 F P Q N A T A C D N A P E A W D L L H R 38

121 GTGCTGCCGA CATTTATCAT CTCCATCTGT TTCTTCGGCC TCCTAGGGAA CCTTTTGTGTC 180
39 V L P T F I I S I C F F G L L G N L F V 58

181 CTGTTGGTCT TCCTCCTGCC CCGGCGGCAA CTGAACGTGG CAGAAATCTA CCTGGCCAAC 240
59 L L V F L L P R R Q L N V A E I Y L A N 78

241 CTGGCAGCCT CTGATCTGGT GTTTGTCTTG GGCTTGCCCT TCTGGGCAGA GAATATCTGG 300
79 L A A S D L V F V L G L P F W A E N I W 98

301 AACCAGTTTA ACTGGCCTTT CGGAGCCCTC CTCTGCCGTG TCATCAATGG GGTCATCAAG 360
99 N Q F N W P F G A L L C R V I N G V I K 118

361 GCCAATTTGT TCATCAGCAT CTTCTGGGTG GTGGCCATCA GCCAGGACCG CTACCGCGTG 420
119 A N L F I S I F L V V A I S Q D R Y R V 138

421 CTGGTGCACC CTATGGCCAG CGGAAGGCAG CAGCGGCGGA GGCAGGCCCG GGTCACCTGC 480
139 L V H P M A S G R Q Q R R R Q A R V T C 158

481 GTGCTCATCT GGGTTGTGGG GGGCCTCTTG AGCATCCCCA CATTCCTGCT GCGATCCATC 540
159 V L I W V V G G L L S I P T F L L R S I 178

541 CAAGCCGTCC CAGATCTGAA CATCACC GCC TGCATCCTGC TCCTCCCCCA TGAGGCCTGG 600
179 Q A V P D L N I T A C I L L L P H E A W 198

601 CACTTTGCAA GGATTGTGGA GTTAAATATT CTGGGTTTCC TCCTACCACT GGCTGCGATC 660
199 H F A R I V E L N I L G F L L P L A A I 218

661 GTCTTCTTCA ACTACCACAT CCTGGCCTCC CTGCGAACGC GGGAGGAGGT CAGCAGGACA 720
219 V F F N Y H I L A S L R T R E E V S R T 238

721 AGAGTGCGGG GGCCGAAGGA TAGCAAGACC ACAGCGCTGA TCCTCACGCT CGTGGTTGCC 780
239 R V R G P K D S K T T A L I L T L V V A 258

781 TTCCTGGTCT GCTGGGCCCC TTACCACTTC TTTGCCTTCC TGGAATTCTT ATTCCAGGTG 840
259 F L V C W A P Y H F F A F L E F L F Q V 278

841 CAAGCAGTCC GAGGCTGCTT TTGGGAGGAC TTCATTGACC TGGGCTGCA ATTGGCCAAC 900
279 Q A V R G C F W E D F I D L G L Q L A N 298

```

Figure 25B

```
901 TTCTTTGCCT TCACTAACAG CTCCTGAAT CCAGTAATTT ATGTCTTTGT GGGCCGGCTC 960
299 F F A F T N S S L N P V I Y V F V G R L 318

961 TTCAGGACCA AGGTCTGGGA ACTTTATAAA CAATGCACCC CTAAAAGTCT TGCTCCAATA 1020
319 F R T K V W E L Y K Q C T P K S L A P I 338

1021 TCTTCATCCC ATAGGAAAGA AATCTTCCAA CTTTCTGGC GGAATTAAAA CAGCATTGAA 1080
339 S S S H R K E I F Q L F W R N * 353
```

1081 CC 1082

1081 CC 1082

Figure 26A

```

1 CTGTGCATGG CATCATCCTG GCCCCCTCTA GAGCTCCAAT CCTCCAACCA GAGCCAGCTC 60
1      M A S S W P P L E L Q S S N Q S Q L 18

61 TTCCCTCAAA ATGCTACGGC CTGTGACAAT GCTCCAGAAG CCTGGGACCT GCTGCACAGA 120
19 F P Q N A T A C D N A P E A W D L L H R 38

121 GTGCTGCCGA CATTTATCAT CTCCATCTGT TTCTTCGGCC TCCTAGGGAA CCTTTTGTGTC 180
39 V L P T F I I S I C F F G L L G N L F V 58

181 CTGTTGGTCT TCCTCCTGCC CCGGCGGCAA CTGAACGTGG CAGAAATCTA CCTGGCCAAC 240
59 L L V F L L P R R Q L N V A E I Y L A N 78

241 CTGGCAGCCT CTGATCTGGT GTTTGTCTTG GGCTTGCCCT TCTGGGCAGA GAATATCTGG 300
79 L A A S D L V F V L G L P F W A E N I W 98

301 AACCAGTTTA ACTGGCCTTT CGGAGCCCTC CTCTGCCGTG TCATCAACGG GGTCATCAAG 360
99 N Q F N W P F G A L L C R V I N G V I K 118

361 GCCAATTTGT TCATCAGCAT CTTCTCTGGT GTGGCCATCA GCCAGGACCG CTACCGCGTG 420
119 A N L F I S I F L V V A I S Q D R Y R V 138

421 CTGGTGACCC CTATGGCCAG CGGAAGGCAG CAGCGGCGGA GACAGGCCCG GGTCACCTGC 480
139 L V H P M A S G R Q Q R R R Q A R V T C 158

481 GTGCTCATCT GGGTTGTGGG GGGCCTCTTG AGCATCCCCA CATTCCTGCT GCGATCCATC 540
159 V L I W V V G G L L S I P T F L L R S I 178

541 CAAGCCGTCC CAGATCTGAA CATCACCGCC TGCATCCTGC TCCTCCCCCA TGAGGCCTGG 600
179 Q A V P D L N I T A C I L L L P H E A W 198

601 CACTTTGCAA GGATTGTGGA GTTAAATATT CTGGGTTTCC TCCTACCACT GGCTGCGATC 660
199 H F A R I V E L N I L G F L L P L A A I 218

661 GTCTTCTTCA ACTACCACAT CCTGGCCTCC CTGCGAACGC GGGAGGAGGT CAGCAGGACA 720
219 V F F N Y H I L A S L R T R E E V S R T 238

721 AGAGTGCGGG GGCCGAAGGA TAGCAAGACC ACAGCGCTGA TCCTCACGCT CGTGGTTGCC 780
239 R V R G P K D S K T T A L I L T L V V A 258

781 TTCCTGGTCT GCTGGGCCCC TTACCACTTC TTTGCCTTCC TGGAATTCTT ATTCCAGGTG 840
259 F L V C W A P Y H F F A F L E F L F Q V 278

841 CAAGCAGTCC GAGGCTGCTT TTGGGAGGAC TTCATTGACC TGGGCCTGCA ATTGGCCAAC 900
279 Q A V R G C F W E D F I D L G L Q L A N 298

```

Figure 26B

```
901 TTCTTTGCCT TCACTAACAG CTCCCTGAAT CCAGTAATTT ATGTCTTTGT GGGCCGGCTC 960
299 F F A F T N S S L N P V I Y V F V G R L 318

961 TTCAGGACCA AGGTCTGGGA ACTTTATAAA CAATGCACCC CTAAAAGTCT TGCTCCAATA 1020
319 F R T K V W E L Y K Q C T P K S L A P I 338

1021 TCTTCATCCC ATAGGAAAGA AATCTTCCAA CTTTCTGGC GGAATTAAAA CAGCATTGAA 1080
339 S S S H R K E I F Q L F W R N * 353

1081 CC 1082
```

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 27A

```

1 CTGTGCATGG CATCATCCTG GCCCCCTCTA GAGCTCCAAT CCTCCAACCA GAGCCAGCTC 60
1      M A S S W P P L E L Q S S N Q S Q L 18

61 TTCCCTCAAA ATGCTACGGC CTGTGACAAT GCTCCAGAAG CCTGGGACCT GCTGCACAGA 120
19 F P Q N A T A C D N A P E A W D L L H R 38

121 GTGCTGCCGA CATTTATCAT CTCCATCTGT TTCTTCGGCC TCCTAGGGAA CCTTTTGTGTC 180
39 V L P T F I I S I C F F G L L G N L F V 58

181 CTGTTGGTCT TCCTCCTGCC CCGGCGGCAA CTGAACGTGG CAGAAATCTA CCTGGCCAAC 240
59 L L V F L L P R R Q L N V A E I Y L A N 78

241 CTGGCAGCCT CTGATCTGGT GTTTGTCTTG GGCTTGCCCT TCTGGGCAGA GAATATCTGG 300
79 L A A S D L V F V L G L P F W A E N I W 98

301 AACCAGTTTA ACTGGCCTTT CGGAGCCCTC CTCTGCCGTG TCATCAACGG GGTCACTAAG 360
99 N Q F N W P F G A L L C R V I N G V I K 118

361 GCCAATTTGT TCATCAGCAT CTTCTGGTGG GTGGCCATCA GCCAGGACCG CTACCGCGTG 420
119 A N L F I S I F L V V A I S Q D R Y R V 138

421 CTGGTGACCC CTATGGCCAG CGGAAGGCAG CAGCGGCGGA GGCAGGCCCG GGTACCTGTC 480
139 L V H P M A S G R Q Q R R R Q A R V T C 158

481 GTGCTCATCT GGGTTGTGGG GGGCCTCTTG AGCATCCCCA CATTCCTGCT GCGATCCATC 540
159 V L I W V V G G L L S I P T F L L R S I 178

541 CAAGCCGTCC CAGATCTGAA CATCACCGCC TGCATCGTGC TCCTCCCCCA TGAGGCCTGG 600
179 Q A V P D L N I T A C I V L L P H E A W 198

601 CACTTTGCAA GGATTGTGGA GTTAAATATT CTGGGTTTCC TCCTACCACT GGCTGCGATC 660
199 H F A R I V E L N I L G F L L P L A A I 218

661 GTCTTCTTCA ACTACCACAT CCTGGCCTCC CTGCGAACGC GGGAGGAGGT CAGCAGGACA 720
219 V F F N Y H I L A S L R T R E E V S R T 238

721 AGAGTGCGGG GGCCGAAGGA TAGCAAGACC ACAGCGCTGA TCCTCACGCT CGTGGTTGCC 780
239 R V R G P K D S K T T A L I L T L V V A 258

781 TTCCTGGTCT GCTGGGCCCC TTACCACTTC TTTGCCTTCC TGGAATTCTT ATTCCAGGTG 840
259 F L V C W A P Y H F F A F L E F L F Q V 278

841 CAAGCAGTCC GAGGCTGCTT TTGGGAGGAC TTCATTGACC TGGGCCTGCA ATTGGCCAAC 900
279 Q A V R G C F W E D F I D L G L Q L A N 298

```

Figure 27B

901	TTCTTGCCT	TCACTAACAG	CTCCCTGAAT	CCAGTAATTT	ATGTCTTTGT	GGGCCGGCTC	960
299	F F A F	T N S	S L N	P V I Y	V F V	G R L	318
961	TTCAGGACCA	AGGTCTGGGA	ACTTTATAAA	CAATGCACCC	CTAAAAGTCT	TGCTCCAATA	1020
319	F R T K	V W E	L Y K	Q C T P	K S L	A P I	338
1021	TCTTCATCCC	ATAGGAAAGA	AATCTTCCAA	CTTTTCTGGC	GGAATTAAAA	CAGCATTGAA	1080
339	S S S H	R K E	I F Q	L F W R	N *		353
1081	CC	1082					

Figure 28A

```

1 CTGTGCATGG CATCATCCTG GCCCCCTCTA GAGCTCCAAT CCTCCAACCA GAGCCAGCTC 60
1      M A S S W P P L E L Q S S N Q S Q L 18

61 TTCCCTCAAA ATGCTACGGC CTGTGACAAT GCTCCAGAAG CCTGGGACCT GCTGCACAGA 120
19 F P Q N A T A C D N A P E A W D L L H R 38

121 GTGCTGCCGA CATTATCAT CTCCATCTGT TTCTTCGGCC TCCTAGGGAA CCTTTTGTGTC 180
39 V L P T F I I S I C F F G L L G N L F V 58

181 CTGTTGGTCT TCCTCCTGCC CCGGCGGCAA CTGAACGTGG CAGAAATCTA CCTGGCCAAC 240
59 L L V F L L P R R Q L N V A E I Y L A N 78

241 CTGGCAGCCT CTGATCTGGT GTTTGTCTTG GGCTTGCCCT TCTGGGCAGA GAATATCTGG 300
79 L A A S D L V F V L G L P F W A E N I W 98

301 AACCAGTTTA ACTGGCCTTT CGGAGCCCTC CTCTGCCGTG TCATCAACGG GGTCATCAAG 360
99 N Q F N W P F G A L L C R V I N G V I K 118

361 GCCAATTTGT TCATCAGCAT CTTCTGGGTG GTGGCCATCA GCCAGGACCG CTACCGCGTG 420
119 A N L F I S I F L V V A I S Q D R Y R V 138

421 CTGGTGACCC CTATGGCCAG CGGAAGGCAG CAGCGGCGGA GGCAGGCCCG GGTCACCTGC 480
139 L V H P M A S G R Q Q R R R Q A R V T C 158

481 GTGCTCATCT GGGTTGTGGG GGGCCTCTTG AGCATCCCCA CATTCCTGCT GCGATCCATC 540
159 V L I W V V G G L L S I P T F L L R S I 178

541 CAAGCCGTCC CAGATCTGAA CATCACCGCC TGCATCCTGC TCCTCCCCCA TGAGGCCTGG 600
179 Q A V P D L N I T A C I L L L P H E A W 198

601 CACTTTGCAA GGATTGTGGA GTTAAATATT CTGGGTTTCC TCCTACCACT GGCTGCGATC 660
199 H F A R I V E L N I L G F L L P L A A I 218

661 GTCTTCTTCA ACTACCACAT CCTGGCCTCC CTGCGAACGC GGGAGGAGGT CAGCAGGACA 720
219 V F F N Y H I L A S L R T R K E V S R T 238

721 AGAGTGCGGG GGCCGAAGGA TAGCAAGACC ACAGCGCTGA TCCTCACGCT CGTGGTTGCC 780
239 R V R G P K D S K T T A L I L T L V V A 258

781 TTCCTGGTCT GCTGGGCCCC TTACCACTTC TTTGCCTTCC TGAATTCTT ATTCCAGGTG 840
259 F L V C W A P Y H F F A F L E F L F Q V 278

841 CAAGCAGTCC GAGGCTGCCT TTGGGAGGAC TTCATTGACC TGGGCCTGCA ATTGGCCAAC 900
279 Q A V R G C F W E D F I D L G L Q L A N 298

```

Figure 28B

901 TTCTTTGCCT TCACTAACAG CTCCCTGAAT CCAGTAATTT ATGTCTTTGT GGGCCGGCTC 960
299 F F A F T N S S L N P V I Y V F V G R L 318

961 TTCAGGACCA AGGTCTGGGA ACTTTATAAA CAATGCACCC CTAAAAGTCT TGCTCCAATA 1020
319 F R T K V W E L Y K Q C T P K S L A P I 338

1021 TCTTCATCCC ATAGGAAAGA AATCTTCCAA CTTTCTGGC GGAATTAAAA CAGCATTGAA 1080
339 S S S H R K E I F Q L F W R N * 353

1081 CC 1082

1081 CC 1082

Figure 29A

```

1  ATGTTCTCTC CCTGGAAGAT ATCAATGTTT CTGTCTGTTT GTGAGGACTC CGTGCCCACC 60
1  M  F  S  P   W  K  I   S  M  F   L  S  V  C   E  D  S   V  P  T   20

61  ACGGCCTCTT TCAGCGCCGA CATGCTCAAT GTCACCTTGC AAGGGCCAC TCTTAACGGG 120
21  T  A  S  F   S  A  D   M  L  N   V  T  L  Q   G  P  T   L  N  G   40

121 ACCTTTGCCC AGAGCAAATG CCCCCAAGTG GAGTGGCTGG GCTGGCTCAA CACCATCCAG 180
41  T  F  A  Q   S  K  C   P  Q  V   E  W  L  G   W  L  N   T  I  Q   60

181 CCCCCCTTCC TCTGGGTGCT GTTCGTGCTG GCCACCCTAG AGAACATCTT TGTCTCAGC 240
61  P  P  F  L   W  V  L   F  V  L   A  T  L  E   N  I  F   V  L  S   80

241 GTCTTCTGCC TGCACAAGAG CAGCTGCACG GTGGCAGAGA TCTACCTGGG GAACCTGGCC 300
81  V  F  C  L   H  K  S   S  C  T   V  A  E  I   Y  L  G   N  L  A   100

301 GCAGCAGACC TGATCCTGGC CTGCGGGCTG CCCTTCTGGG CCATCACCAT CTCCAACAAC 360
101 A  A  D  L   I  L  A   C  G  L   P  F  W  A   I  T  I   S  N  N   120

361 TTCGACTGGC TCTTTGGGGA GACGCTCTGC CGCGTGGTGA ATGCCATTAT CTCCATGAAC 420
121 F  D  W  L   F  G  E   T  L  C   R  V  V  N   A  I  I   S  M  N   140

421 CTGTACAGCA GCATCTGTTT CCTGATGCTG GTGAGCATCG ACCGCTACCT GGCCCTGGTG 480
141 L  Y  S  S   I  C  F   L  M  L   V  S  I  D   R  Y  L   A  L  V   160

481 AAAACCATGT CCATGGGCCG GATGCGCGGC GTGCGCTGGG CCAAGCTCTA CAGCTTGGTG 540
161 K  T  M  S   M  G  R   M  R  G   V  R  W  A   K  L  Y   S  L  V   180

541 ATCTGGGGGT GTACGCTGCT CCTGAGCTCA CCCATGCTGG TGTCCGGAC CATGAAGGAG 600
181 I  W  G  C   T  L  L   L  S  S   P  M  L  V   F  R  T   M  K  E   200

601 TACAGCGATG AGGGCCACAA CGTCACCGCT TGTGTCATCA GCTACCCATC CCTCATCTGG 660
201 Y  S  D  E   G  H  N   V  T  A   C  V  I  S   Y  P  S   L  I  W   220

661 GAAGTGTTCA CCAACATGCT CCTGAATGTC GTGGGCTTCC TGCTGCCCCT GAGTGTCAATC 720
221 E  V  F  T   N  M  L   L  N  V   V  G  F  L   L  P  L   S  V  I   240

721 ACCTTCTGCA CGATGCAGAT CATGCAGGTG CTGCGGAACA ACGAGATGCA GAAGTTCAAG 780
241 T  F  C  T   M  Q  I   M  Q  V   L  R  N  N   E  M  Q   K  F  K   260

781 GAGATCCAGA CGGAGAGGAG GGCCACGGTG CTAGTCCTGG TTGTGCTGCT GCTATTCATC 840
261 E  I  Q  T   E  R  R   A  T  V   L  V  L  V   V  L  L   L  F  I   280

841 ATCTGCTGGC TGCCCTTCCA GATCAGCACC TTCCTGGATA CGCTGCATCG CCTCGGCATC 900
281 I  C  W  L   P  F  Q   I  S  T   F  L  D  T   L  H  R   L  G  I   300

```

Figure 29B

```

901 CTCTCCAGCT GCCAGGACGA GGCATCATC GATGTAATCA CACAGATCGC CTCCTTCATG 960
301 L S S C Q D E R I I D V I T Q I A S F M 320

961 GCCTACAGCA ACAGCTGCCT CAACCCACTG GTGTACGTGA TCGTGGGCAA GCGCTTCCGA 1020
321 A Y S N S C L N P L V Y V I V G K R F R 340

1021 AAGAAGTCTT GGGAGGTGTA CCAGGGAGTG TGCCAGAAAG GGGGCTGCAG GTCAGAACCC 1080
341 K K S W E V Y Q G V C Q K G G C R S E P 360

1081 ATTACAGATGG AGAACTCCAT GGGCACACTG CGGACCTCCA TCTCCGTGGA ACGCCAGATT 1140
361 I Q M E N S M G T L R T S I S V E R Q I 380

1141 CACAACTGC AGGACTGGGC AGGGAGCAGA CAGTGAGCAA ACGCCAGCAG GGCTGCTGTG 1200
381 H K L Q D W A G S R Q * 391

1201 AATTTGTGTA AGGATTGAGG GACAGTTGCT TTTCAGCATG GGCCAGGAA TGCCAAGGAG 1260

1261 ACATCTATGC ACGACCTTGG GAAATGAGTT GATGTCTCCG GTAAAACACC GGAGACTAAT 1320

1321 TCCTGCCCTG CCCAATTTTG CAGGGAGCAT GGCTGTGAGG ATGGGGTGAA CTCACGCACA 1380

1381 GCCAAGGACT CCAAATCAC AACAGCATT CTGTTCTTAT TTGCTGCCAC ACCTGAGCCA 1440

1441 GCCTGCTCCT TCCCAGGAGT GGAGGAGGCC TGGGGGAGG GAGAGGAGTG ACTGAGCTTC 1500

1501 CCTCCCGTGT GTTCTCCGTC CCTGCCCCAG CAAGACAACT TAGATCTCCA GGAGAACTGC 1560

1561 CATCCAGCTT TGGTGCAATG GCTGAGTGCA CAAGTGAGTT GTTGCCCTGG GTTCTTTTAA 1620

1621 TCTATTCAGC TAGAACTTTG AAGGACAATT TCTTGCAATTA ATAAAGGTTA AGCCCTGAGG 1680

1681 GGTCCTGAT AACAACTGG AGACCAGGAT TTTATGGCTC CCCTCACTGA TGGACAAGGA 1740

1741 GGTCTGTGCC AAAGAAGAAT CCAATAAGCA CATATTGAGC ACTTGCTGTA TATGCAGTAT 1800

1801 TGAGCACTGT AGGCAAGACC CAAGAAAGAG AAGGAGCCAT CTCCATCTTG AAGGAACTCA 1860

1861 AAGACTCAAG TGGGAACGAC TGGGCACTGC CACCACCAGA AAGCTGTTTC ACGAGACGGT 1920

1921 CGAGCAGGGT GCTGTGGGTG ATATGGACAG CAGAAGGGGG AGACCAAGGT TCCAGCTCAA 1980

1981 CCAATAACTA TTGCACAACC ACCTGTCCCT GCCTCAGTTC CCTTTTATGT AACATGAAGT 2040

```

Figure 29C

```

2041 CGTTGTGAGG GTTAAAGGCA GTAACAGGTA TAAAGTACTT AGAAAAGCAA AGGGTGCTAC 2100
2101 GTACATGTGA GGCATCATTÀ CGCAGACGTA ACTGGGATAT GTTTACTATA AGGAAAAGAC 2160
2161 ACTGAGGTCT AGAAATAGCT CCGTGGAGCA GAATCAGTAT TGGGAGCCGG TGGCGGTGTG 2220
2221 AAGCACCAGT GTCTGGCACA CAGTAGGTGC TCATTGGCTC CCTTCCACCT GTCATTCCCA 2280
2281 CCACCCTGAG GCCCCAACCG CCACACACAC AGGAGCATTT GGAGAGAAGG CCATGTCTTC 2340
2341 AAAGTCTGAT TTGTGATGAG GCAGAGGAAG ATATTCTAA TCGGTCTTGC CCAGAGGATC 2400
2401 ACAGTGCTGA GACCCCCCAC CACCAGCCGG TACCTGGGAA GGGGAGAGT GCAGGCCTGC 2460
2461 TCAGGGACTG TTCCTGTCTC AGCAACCAAG GGATTGTTC TGTCAATCAA TGTTTATTG 2520
2521 GAAGGTGGCC CAGTATGAGC CCTAGAAGAG TGTGAAAAGG AATGGCAATG GTGTTACCA 2580
2581 TCGGCAGTGC CAGGGCAGCA CTCATTCACT TGATAAATGA ATATTTATTA GCTGGTTGGA 2640
2641 GAGCTAGAAC CTGGAGAGCT AGAACCTGGA GAACTAGAAC CTGGAGGGCT AGAACCTGGA 2700
2701 GAGGCTAGAA CCAAGAAGGG CTAGAACCTG GAGGGGCTAG AACCTAGAGA AGCTAAAACC 2760
2761 TGAGCTAGAA GCTGGAGGAC TAGAACCTGG AGGGCTGGAA TCTGAAGGGC TAGAACCTGG 2820
2821 AGGGCTGGAA TCTGGAGAGC TAGAACCTGG AGGGCTAGAA CCTGGAGGGC TAGAACCTAG 2880
2881 AAGGGCTAGA ACCTGGAGGG CTGGAATCTG GAGAGCTAGA ACCTGGAGGG CTAGAACCTG 2940
2941 GAGGGCTAGA ACCTAGAAGG GCTAGAACCT GGAGGGCTAG AACCTGGCAG GTTAGAACCT 3000
3001 AGAAGGGCTA GAACCTGGAG AGCCAGAACC TGGAGGGCTA GAACCTGGAA GGGCTAGAAC 3060
3061 CTGTAGAGCT AGAACATGGA GAGCTAGAAC CCGGCAGGCT AGAACCTGGC AAGCTAGAAC 3120
3121 CTGGAGGGAA TGAACCTGGA GGGCTAGAAC CTGGAGAATG AGAAAAATTT ACATGGCAAA 3180
3181 GAGCCCATAA ATCCTGACCA ATCCAACCT GAATTTTAAA GCAAAAAGCGT GAAAAAAAAG 3240

```

Figure 29D

3241 ATTCCCTCCT TACCCCCAAC CCACTCTTTT TTCCCACCAC CCACTCTCCT CTGCCTCAGT 3300
3301 AAGTATCTGG AGGAAGAAAA CAGGTGAAAG AAGAAGTAAA AACCATTAG TATTAGTATT 3360
3361 AGAATGAAGT CAAACTGTGC CACACATGGT GAATGAAAAA AAAAAAAAAG AGGCTGTGTT 3420
3421 TTGTCACACA GGGCAGTCAT TCAGCACCAG AGCACGTGAT GGTCTGAGAC TCTCTTAGGA 3480
3481 GCAGAGCTCT GCCGCAATGG CCATGTGGGG ATCCACACCT GGTCTGAGGG GCAACTGAGT 3540
3541 CTGCGGGAGA AGAGCGGCC TATGCATGGT GTAGATGCCC TGATAAAGAA CATCTGTCCT 3600
3601 GTGAAAGACT CAATGAGCTG TTATGTTGTA AACAGGAAGC ATTTACATC CAAACGAGAA 3660
3661 AATCATGTAA ACATGTGTCT TTTCTGTAGA GCATAATAAA TGGATGAGGT TTTTGCAAAA 3720
3721 AAAAAAAAAA AAA 3733

Figure 30A

```

1  ATGTTCTCTC  CCTGGAAGAT  ATCAATGTTT  CTGTCTGTTC  GTGAGGACTC  CGTGCCCACC  60
1  M  F  S  P    W  K  I    S  M  F    L  S  V  R    E  D  S    V  P  T    20

61  ACGGCCTCTT  TCAGCGCCGA  CATGCTCAAT  GTCACCTTGC  AAGGGCCCAC  TCTTAACGGG  120
21  T  A  S  F    S  A  D    M  L  N    V  T  L  Q    G  P  T    L  N  G    40

121  ACCTTTGCCC  AGAGCAAATG  CCCCCAAGTG  GAGTGGCTGG  GCTGGCTCAA  CACCATCCAG  180
41  T  F  A  Q    S  K  C    P  Q  V    E  W  L  G    W  L  N    T  I  Q    60

181  CCCCCCTTCC  TCTGGGTGCT  GTTCGTGCTG  GCCACCCTAG  AGAACATCTT  TGTCTCAGC  240
61  P  P  F  L    W  V  L    F  V  L    A  T  L  E    N  I  F    V  L  S    80

241  GTCTTCTGCC  TGCACAAGAG  CAGCTGCACG  GTGGCAGAGA  TCTACCTGGG  GAACCTGGCC  300
81  V  F  C  L    H  K  S    S  C  T    V  A  E  I    Y  L  G    N  L  A    100

301  GCAGCAGACC  TGATCCTGGC  CTGCGGGCTG  CCCTTCTGGG  CCATCACCAT  CTCCAACAAC  360
101  A  A  D  L    I  L  A    C  G  L    P  F  W  A    I  T  I    S  N  N    120

361  TTCGACTGGC  TCTTTGGGGA  GACGCTCTGC  CGCGTGGTGA  ATGCCATTAT  CTCCATGAAC  420
121  F  D  W  L    F  G  E    T  L  C    R  V  V  N    A  I  I    S  M  N    140

421  CTGTACAGCA  GCATCTGTTT  CCTGATGCTG  GTGAGCATCG  ACCGCTACCT  GGCCCTGGTG  480
141  L  Y  S  S    I  C  F    L  M  L    V  S  I  D    R  Y  L    A  L  V    160

481  AAAACCATGT  CCATGGGCCG  GATGCGCGGC  GTGCGCTGGG  CCAAGCTCTA  CAGCTTGGTG  540
161  K  T  M  S    M  G  R    M  R  G    V  R  W  A    K  L  Y    S  L  V    180

541  ATCTGGGGGT  GTACGCTGCT  CCTGAGCTCA  CCCATGCTGG  TGTTCGGAC  CATGAAGGAG  600
181  I  W  G  C    T  L  L    L  S  S    P  M  L  V    F  R  T    M  K  E    200

601  TACAGCGATG  AGGGCCACAA  CGTCACCGCT  TGTGTCATCA  GCTACCCATC  CCTCATCTGG  660
201  Y  S  D  E    G  H  N    V  T  A    C  V  I  S    Y  P  S    L  I  W    220

661  GAAGTGTTCA  CCAACATGCT  CCTGAATGTC  GTGGGCTTCC  TGCTGCCCCT  GAGTGTCAATC  720
221  E  V  F  T    N  M  L    L  N  V    V  G  F  L    L  P  L    S  V  I    240

721  ACCTTCTGCA  CGATGCAGAT  CATGCAGGTG  CTGCGGAACA  ACGAGATGCA  GAAGTTCAAG  780
241  T  F  C  T    M  Q  I    M  Q  V    L  R  N  N    E  M  Q    K  F  K    260

781  GAGATCCAGA  CGGAGAGGAG  GGCCACGGTG  CTAGTCCTGG  TTGTGCTGCT  GCTATTCATC  840
261  E  I  Q  T    E  R  R    A  T  V    L  V  L  V    V  L  L    L  F  I    280

841  ATCTGCTGGC  TGCCCTTCCA  GATCAGCACC  TTCCTGGATA  CGCTGCATCG  CCTCGGCATC  900
281  I  C  W  L    P  F  Q    I  S  T    F  L  D  T    L  H  R    L  G  I    300

```

Figure 30B

```

901 CTCTCCAGCT GCCAGGACGA GCGCATCATC GACGTAATCA CACAGATCGC CTCCTTCATG 960
301 L S S C Q D E R I I D V I T Q I A S F M 320

961 GCCTACAGCA ACAGCTGCCT CAACCCACTG GTGTACGTGA TCGTGGGCAA GCGCTTCCGA 1020
321 A Y S N S C L N P L V Y V I V G K R F R 340

1021 AAGAAGTCTT GGGAGGTGTG CCAGGGAGTG TGCCAGAAAG GGGGCTGCAG GTCAGAACCC 1080
341 K K S W E V Y Q G V C Q K G G C R S E P 360

1081 ATTCAGATGG AGAACTCCAT GGGCACACTG CGGACCTCCA TCTCCGTGGA ACGCCAGATT 1140
361 I Q M E N S M G T L R T S I S V E R Q I 380

1141 CACAACTGCG AGGACTGGGC AGGGAGCAGA CAGTGAGCAA ACGCCAGCAG GGCTGCTGTG 1200
381 H K L Q D W A G S R Q * 391

1201 AATTTGTGTA AGGATTGAGG GACAGTTGCT TTTCAGCATG GGCCAGGAA TGCCAAGGAG 1260

1261 ACATCTATGC ACGACCTTGG GAAATGAGTT GATGTCTCCG GTAAAACACC GGAGACTAAT 1320

1321 TCCTGCCCTG CCCAATTTTG CAGGGAGCAT GGCTGTGAGG ATGGGGTGAA CTCACGCACA 1380

1381 GCCAAGGACT CCAAATCAC AACAGCATTA CTGTTCTTAT TTGCTGCCAC ACCTGAGCCA 1440

1441 GCCTGCTCCT TCCCAGGAGT GGAGGAGGCC TGGGGGGAGG GAGAGGAGTG ACTGAGCTTC 1500

1501 CCTCCCGTGT GTTCTCCGTC CCTGCCCCAG CAAGACAAC TAGATCTCCA GGAGAACTGC 1560

1561 CATCCAGCTT TGGTGCAATG GCTGAGTGCA CAAGTGAGTT GTTGCCCTGG GTTCTTTTAA 1620

1621 TCTATTCAGC TAGAACTTTG AAGGACAATT TCTTGCATTA ATAAAGGTTA AGCCCTGAGG 1680

1681 GTTCCCTGAT AACAACTTGG AGACCAGGAT TTTATGGCTC CCCTCACTGA TGGACAAGGA 1740

1741 GGTCTGTGCC AAAGAAGAAT CCAATAAGCA CATATTGAGC ACTTGCTGTA TATGCAGTAT 1800

1801 TGAGCACTGT AGGCAAGACC CAAGAAAGAG AAGGAGCCAT CTCCATCTTG AAGGAACTCA 1860

1861 AAGACTCAAG TGGGAACGAC TGGGCACTGC CACCACCAGA AAGCTGTTTC ACGAGACGGT 1920

1921 CGAGCAGGGT GCTGTGGGTG ATATGGACAG CAGAAGGGGG AGACCAAGGT TCCAGCTCAA 1980

1981 CCAATAACTA TTGCACAACC ACCTGTCCCT GCCTCAGTTC CCTTTTATGT AACATGAAGT 2040

```

Figure 30C

2041 CGTTGTGAGG GTTAAAGGCA GTAACAGGTA TAAAGTACTT AGAAAAGCAA AGGGTGCTAC 2100
 2101 GTACATGTGA GGCATCATTA CGCAGACGTA ACTGGGATAT GTTTACTATA AGGAAAAGAC 2160
 2161 ACTGAGGTCT AGAAATAGCT CCGTGGAGCA GAATCAGTAT TGGGAGCCGG TGGCGGTGTG 2220
 2221 AAGCACCAGT GTCTGGCACA CAGTAGGTGC TCATTGGCTC CCTTCCACCT GTCATTCCCA 2280
 2281 CCACCCTGAG GCCCCAACCG CCACACACAC AGGAGCATTT GGAGAGAAGG CCATGTCTTC 2340
 2341 AAAGTCTGAT TTGTGATGAG GCAGAGGAAG ATATTTCTAA TCGGTCTTGC CCAGAGGATC 2400
 2401 ACAGTGCTGA GACCCCCCAC CACCAGCCGG TACCTGGGAA GGGGGAGAGT GCAGGCCTGC 2460
 2461 TCAGGGACTG TTCCTGTCTC AGCAACCAAG GGATTGTTC TGTCAATCAA TGGTTTATTG 2520
 2521 GAAGGTGGCC CAGTATGAGC CCTAGAAGAG TGTGAAAAGG AATGGCAATG GTGTTTACCA 2580
 2581 TCGGCAGTGC CAGGGCAGCA CTCATTCACT TGATAAATGA ATATTTATTA GCTGGTTGGA 2640
 2641 GAGCTAGAAC CTGGAGAGCT AGAACCTGGA GAACTAGAAC CTGGAGGGCT AGAACCTGGA 2700
 2701 GAGGCTAGAA CCAAGAAGGG CTAGAACCTG GAGGGGCTAG AACCTAGAGA AGCTAAAACC 2760
 2761 TGAGCTAGAA GCTGGAGGAC TAGAACCTGG AGGGCTGGAA TCTGAAGGGC TAGAACCTGG 2820
 2821 AGGGCTGGAA TCTGGAGAGC TAGAACCTGG AGGGCTAGAA CCTGGAGGGC TAGAACCTAG 2880
 2881 AAGGGCTAGA ACCTGGAGGG CTGGAATCTG GAGAGCTAGA ACCTGGAGGG CTAGAACCTG 2940
 2941 GAGGGCTAGA ACCTAGAAGG GCTAGAACCT GGAGGGCTAG AACCTGGCAG GTTAGAACCT 3000
 3001 AGAAGGGCTA GAACCTGGAG AGCCAGAACC TGGAGGGCTA GAACCTGGAA GGGCTAGAAC 3060
 3061 CTGTAGAGCT AGAACATGGA GAGCTAGAAC CCGGCAGGCT AGAACCTGGC AAGCTAGAAC 3120
 3121 CTGGAGGGAA TGAACCTGGA GGGCTAGAAC CTGGAGAATG AGAAAAATTT ACATGGCAAA 3180
 3181 GAGCCCATAA ATCTTGACCA ATCCAACCTT GAATTTTAAA GCAAAAGCGT GAAAAAAAAG 3240

Figure 30D

3241	ATTCCTCCT	TACCCCCAAC	CCACTCTTTT	TTCCCACCAC	CCACTCTCCT	CTGCCTCAGT	3300
3301	AAGTATCTGG	AGGAAGAAAA	CAGGTGAAAG	AAGAAGTAAA	AACCATTTAG	TATTAGTATT	3360
3361	AGAATGAAGT	CAAACGTGTC	CACACATGGT	GAATGAAAAA	AAAAAAAAAG	AGGCTGTGTT	3420
3421	TTGTCACACA	GGGCAGTCAT	TCAGCACCAG	AGCACGTGAT	GGTCTGAGAC	TCTCTTAGGA	3480
3481	GCAGAGCTCT	GCCGCAATGG	CCATGTGGGG	ATCCACACCT	GGTCTGAGGG	GCAACTGAGT	3540
3541	CTGCGGGAGA	AGAGCGGCCC	TATGCATGGT	GTAGATGCCC	TGATAAAGAA	CATCTGTCCT	3600
3601	GTGAAAGACT	CAATGAGCTG	TTATGTTGTA	AACAGGAAGC	ATTTCACATC	CAAACGAGAA	3660
3661	AATCATGTAA	ACATGTGTCT	TTTCTGTAGA	GCATAATAAA	TGGATGAGGT	TTTGTCAAAA	3720
3721	AAAAAAAAAA	AAA	3733				

Figure 31A

```

1  ATGTTCTCTC  CCTGGAAGAT  ATCAATGTTT  CTGTCTGTTT  GTGAGGACTC  CGTGCCCACC  60
1  M  F  S  P    W  K  I    S  M  F    L  S  V  R    E  D  S    V  P  T    20

61  ACGGCCTCTT  TCAGCGCCGA  CATGCTCAAT  GTCACCTTGC  AAGGGCCAC  TCTTAACGGG  120
21  T  A  S  F    S  A  D    M  L  N    V  T  L  Q    G  P  T    L  N  G    40

121  ACCTTTGCCC  AGAGCAAATG  CCCCCAAGTG  GAGTGGCTGG  GCTGGCTCAA  CACCATCCAG  180
41  T  F  A  Q    S  K  C    P  Q  V    E  W  L  G    W  L  N    T  I  Q    60

181  CCCCCCTTCC  TCTGGGTGCT  GTTCGTGCTG  GCCACCCTAG  AGAACATCTT  TGTCTCAGC  240
61  P  P  F  L    W  V  L    F  V  L    A  T  L  E    N  I  F    V  L  S    80

241  GTCTTCTGCC  TGCACAAGAG  CAGCTGCACG  GTGGCAGAGA  TCTACCTGGG  GAACCTGGCC  300
81  V  F  C  L    H  K  S    S  C  T    V  A  E  I    Y  L  G    N  L  A    100

301  GCAGCAGACC  TGATCCTGGC  CTGCGGGCTG  CCCTTCTGGG  CCATCACCAT  CTCCAACAAC  360
101  A  A  D  L    I  L  A    C  G  L    P  F  W  A    I  T  I    S  N  N    120

361  TTCGACTGGC  TCTTTGGGGA  GACGCTCTGC  CGCGTGGTGA  ATGCCATTAT  CTCCATGAAC  420
121  F  D  W  L    F  G  E    T  L  C    R  V  V  N    A  I  I    S  M  N    140

421  CTGTACAGCA  GCATCTGTTT  CCTGATGCTG  GTGAGCATCG  ACCGCTACCT  GGCCCTGGTG  480
141  L  Y  S  S    I  C  F    L  M  L    V  S  I  D    R  Y  L    A  L  V    160

481  AAAACCATGT  CCATGGGCCG  GATGCGCGGC  GTGCGCTGGG  CCAAGCTCTA  CAGCTTGGTG  540
161  K  T  M  S    M  G  R    M  R  G    V  R  W  A    K  L  Y    S  L  V    180

541  ATCTGGGGGT  GTACGCTGCT  CCTGAGCTCA  CCCATGCTGG  TGTTCGGAC  CATGAAGGAG  600
181  I  W  G  C    T  L  L    L  S  S    P  M  L  V    F  R  T    M  K  E    200

601  TACAGCGATG  AGGGCCACAA  CGTCACCGCT  TGTGTCATCA  GCTACCCATC  CCTCATCTGG  660
201  Y  S  D  E    G  H  N    V  T  A    C  V  I  S    Y  P  S    L  I  W    220

661  GAAGTGTTCA  CCAACATGCT  CCTGAATGTC  GTGGGCTTCC  TGCTGCCCTT  GAGTGTATC  720
221  E  V  F  T    N  M  L    L  N  V    V  G  F  L    L  P  L    S  V  I    240

721  ACCTTCTGCA  CGATGCAGAT  CATGCAGGTG  CTGCGGAACA  ACGAGATGCA  GAAGTTCAAG  780
241  T  F  C  T    M  Q  I    M  Q  V    L  R  N  N    E  M  Q    K  F  K    260

781  GAGATCCAGA  CGGAGAGGAG  GGCCACGGTG  CTAGTCCTGG  TTGTGCTGCT  GCTATTCATC  840
261  E  I  Q  T    E  R  R    A  T  V    L  V  L  V    V  L  L    L  F  I    280

841  ATCTGCTGGC  TGCCCTTCCA  GATCAGCACC  TTCCTGGATA  CGCTGCATCG  CCTCGGCATC  900
281  I  C  W  L    P  F  Q    I  S  T    F  L  D  T    L  H  R    L  G  I    300

```

Figure 31B

```

901 CTCTCCAGCT GCCAGGACGA GCGCATCATC GATGTAATCA CACAGATCGC CTCCTTCATG 960
301 L S S C Q D E R I I D V I T Q I A S F M 320

961 GCCTACAGCA ACAGCTGCCT CAACCCACTG GTGTACGTGA TCGTGGGCAA GCGCTTCCGA 1020
321 A Y S N S C L N P L V Y V I V G K R F R 340

1021 AAGAAGTCTT GGGAGGTGTA CCAGGGAGTG TGCCAGAAAG AGGGCTGCAG GTCAGAACCC 1080
341 K K S W E V Y Q G V C Q K E G C R S E P 360

1081 ATTCAGATGG AGAACTCCAT GGGCACACTG CGGACCTCCA TCTCCGTGGA ACGCCAGATT 1140
361 I Q M E N S M G T L R T S I S V E R Q I 380

1141 CACAAACTGC AGGACTGGGC AGGGAGCAGA CAGTGAGCAA ACGCCAGCAG GGCTGCTGTG 1200
381 H K L Q D W A G S R Q * 391

1201 AATTTGTGTA AGGATTGAGG GACAGTTGCT TTTCAGCATG GGCCCAGGAA TGCCAAGGAG 1260

1261 ACATCTATGC ACGACCTTGG GAAATGAGTT GATGTCTCCG GTAAAACACC GGAGACTAAT 1320

1321 TCCTGCCCTG CCCAATTTTG CAGGGAGCAT GGCTGTGAGG ATGGGGTGAA CTCACGCACA 1380

1381 GCCAAGGACT CCAAATCAC AACAGCATTA CTGTTCTTAT TTGCTGCCAC ACCTGAGCCA 1440

1441 GCCTGCTCCT TCCCAGGAGT GGAGGAGGCC TGGGGGAGG GAGAGGAGTG ACTGAGCTTC 1500

1501 CCTCCCGTGT GTTCTCCGTC CCTGCCCCAG CAAGACAACT TAGATCTCCA GGAGAACTGC 1560

1561 CATCCAGCTT TGGTGCAATG GCTGAGTGCA CAAGTGAGTT GTTGCCCTGG GTTCTTTTAA 1620

1621 TCTATTCAGC TAGAACTTTG AAGGACAATT TCTTGCATTA ATAAAGGTTA AGCCCTGAGG 1680

1681 GGTCCCTGAT AACAACTTGG AGACCAGGAT TTTATGGCTC CCCTCACTGA TGGACAAGGA 1740

1741 GGTCTGTGCC AAAGAAGAAT CCAATAAGCA CATATTGAGC ACTTGCTGTA TATGCAGTAT 1800

1801 TGAGCACTGT AGGCAAGACC CAAGAAAGAG AAGGAGCCAT CTCCATCTTG AAGGAACTCA 1860

1861 AAGACTCAAG TGGGAACGAC TGGGCACTGC CACCACCAGA AAGCTGTTCG ACGAGACGGT 1920

1921 CGAGCAGGGT GCTGTGGGTG ATATGGACAG CAGAAGGGGG AGACCAAGGT TCCAGCTCAA 1980

1981 CCAATAACTA TTGCACAACC ACCTGTCCCT GCCTCAGTTC CCTTTTATGT AACATGAAGT 2040

```

Figure 31C

2041 CGTTGTGAGG GTTAAAGGCA GTAACAGGTA TAAAGTACTT AGAAAAGCAA AGGGTGCTAC 2100
 2101 GTACATGTGA GGCATCATTA CGCAGACGTA ACTGGGATAT GTTTACTATA AGGAAAAGAC 2160
 2161 ACTGAGGTCT AGAAATAGCT CCGTGGAGCA GAATCAGTAT TGGGAGCCGG TGGCGGTGTG 2220
 2221 AAGCACCAGT GTCTGGCACA CAGTAGGTGC TCATTGGCTC CCTTCCACCT GTCATTCCCA 2280
 2281 CCACCCTGAG GCCCCAACCG CCACACACAC AGGAGCATTT GGAGAGAAGG CCATGTCTTC 2340
 2341 AAAGTCTGAT TTGTGATGAG GCAGAGGAAG ATATTCTAA TCGGTCTTGC CCAGAGGATC 2400
 2401 ACAGTGCTGA GACCCCCCAC CACCAGCCGG TACCTGGGAA GGGGGAGAGT GCAGGCCTGC 2460
 2461 TCAGGGACTG TTCCTGTCTC AGCAACCAAG GGATTGTTC TGTCAATCAA TGGTTTATTG 2520
 2521 GAAGGTGGCC CAGTATGAGC CCTAGAAGAG TGTGAAAAGG AATGGCAATG GTGTTCACCA 2580
 2581 TCGGCAGTGC CAGGGCAGCA CTCATTCACT TGATAAATGA ATATTTATTA GCTGGTTGGA 2640
 2641 GAGCTAGAAC CTGGAGAGCT AGAACCTGGA GAACTAGAAC CTGGAGGGCT AGAACCTGGA 2700
 2701 GAGGCTAGAA CCAAGAAGGG CTAGAACCTG GAGGGGCTAG AACCTAGAGA AGCTAAAACC 2760
 2761 TGAGCTAGAA GCTGGAGGAC TAGAACCTGG AGGGCTGGAA TCTGAAGGGC TAGAACCTGG 2820
 2821 AGGGCTGGAA TCTGGAGAGC TAGAACCTGG AGGGCTAGAA CCTGGAGGGC TAGAACCTAG 2880
 2881 AAGGGCTAGA ACCTGGAGGG CTGGAATCTG GAGAGCTAGA ACCTGGAGGG CTAGAACCTG 2940
 2941 GAGGGCTAGA ACCTAGAAGG GCTAGAACCT GGAGGGCTAG AACCTGGCAG GTTAGAACCT 3000
 3001 AGAAGGGCTA GAACCTGGAG AGCCAGAACC TGGAGGGCTA GAACCTGGAA GGGCTAGAAC 3060
 3061 CTGTAGAGCT AGAACATGGA GAGCTAGAAC CCGGCAGGCT AGAACCTGGC AAGCTAGAAC 3120
 3121 CTGGAGGGAA TGAACCTGGA GGGCTAGAAC CTGGAGAATG AGAAAAATTT ACATGGCAAA 3180
 3181 GAGCCCATAA ATCTGACCA ATCCAACCT GAATTTTAA GCAAAAGCGT GAAAAAAAAG 3240

Figure 31D

3241 ATTCCCTCCT TACCCCAAC CCACTCTTTT TTCCCACCAC CCACTCTCCT CTGCCTCAGT 3300
3301 AAGTATCTGG AGGAAGAAA CAGGTGAAAG AAGAAGTAA AACCATTAG TATTAGTATT 3360
3361 AGAATGAAGT CAAACTGTGC CACACATGGT GAATGAAAA AAAAAAAG AGGCTGTGTT 3420
3421 TTGTCACAC GGCAGTCAT TCAGCACCAG AGCACGTGAT GGTCTGAGAC TCTCTTAGGA 3480
3481 GCAGAGCTCT GCCGCAATGG CCATGTGGGG ATCCACACCT GGTCTGAGGG GCAACTGAGT 3540
3541 CTGCGGGAGA AGAGCGGCC TATGCATGGT GTAGATGCCC TGATAAAGAA CATCTGTCCT 3600
3601 GTGAAAGACT CAATGAGCTG TTATGTTGTA AACAGGAAGC ATTTACATC CAAACGAGAA 3660
3661 AATCATGTAA ACATGTGTCT TTTCTGTAGA GCATAATAAA TGGATGAGGT TTTTGCAAAA 3720
3721 AAAAAAAAA AAA 3733

Figure 32A

1	CGCCCAACCCAAGTTCAAAGGCTGATAAGAGAGAAAAATCTCATGAGGAGGTTTTAGTCTA	60
61	GGGAAAGTCATTTCAGTGGATGTGATCTTGGCTCACAGGGGACGATGTCAAGCTCTTCCTG	120
1	M S S S S W	6
121	GCTCCTTCTCAGCCTTGTTGCTGTAAGTCTGCTCAGTCCACCATTTAGGAACAGGCCAA	180
7	L L L S L V A V T A A Q S T I E E Q A K	26
181	GACATTTTTGGACAAGTTTAACCACGAAGCCGAAGACCTGTTCTATCAAAGTTCACTTGC	240
27	T F L D K F N H E A E D L F Y Q S S L A	46
241	TTCTTGGAATTATAACACCAATATTACTGAAGAGAATGTCCAAAACATGAATAATGCTGG	300
47	S W N Y N T N I T E E N V Q N M N N A G	66
301	GGACAAATGGTCTGCCTTTTTAAAGGAACAGTCCACACTTGCCCAAATGTATCCACTACA	360
67	D K W S A F L K E Q S T L A Q M Y P L Q	86
361	AGAAATTCAGAATCTCACAGTCAAGCTTCAGCTGCAGGCTCTTCAGCAAAATGGGTCTTC	420
87	E I Q N L T V K L Q L Q A L Q Q N G S S	106
421	AGTGCTCTCAGAAGACAAGAGCAAACGGTTGAACACAATTTCTAAATACAATGAGCACCAT	480
107	V L S E D K S K R L N T I L N T M S T I	126
481	CTACAGTACTGGAAAAGTTTGTAACCCAGATAATCCACAAGAATGCTTATTACTTGAACC	540
127	Y S T G K V C N P D N P Q E C L L L E P	146
541	AGGTTTGAATGAAATAATGGCAAACAGTTTAGACTACAATGAGAGGCTCTGGGCTTGGGA	600
147	G L N E I M A N S L D Y N E R L W A W E	166
601	AAGCTGGAGATCTGAGGTCGGCAAGCAGCTGAGGCCATTATATGAAGAGTATGTGGTCTT	660
167	S W R S E V G K Q L R P L Y E E Y V V L	186
661	GAAAAATGAGATGGCAAGAGCAAATCATTATGAGGACTATGGGGATTATTGGAGAGGAGA	720
187	K N E M A R A N H Y E D Y G D Y W R G D	206
721	CTATGAAGTAAATGGGGTAGATGGCTATGACTACAGCCGCGCCAGTTGATTGAAGATGT	780
207	Y E V N G V D G Y D Y S R G Q L I E D V	226
781	GGAACATACCTTTGAAGAGATTAAACCATTATATGAACATCTTCATGCCTATGTGAGGGC	840
227	E H T F E E I K P L Y E H L H A Y V R A	246
841	AAAGTTGATGAATGCCTATCCTTCCTATATCAGTCCAATTTGGATGCCTCCCTGCTCATTT	900
247	K L M N A Y P S Y I S P I G C L P A H L	266

Downloaded from www.sciencedirect.com

Figure 32B

901	GCTTGGTGATATGTGGGGTAGATTTTGGACAAATCTGTACTCTTTGACAGTTCCCTTTGG	960
267	L G D M W G R F W T N L Y S L T V P F G	286
961	ACAGAAACCAAACATAGATGTTACTGATGCAATGGTGGACCAGGCCTGGGATGCACAGAG	1020
287	Q K P N I D V T D A M V D Q A W D A Q R	306
1021	AATATTCAAGGAGGCCGAGAAGTTCTTTGTATCTGTTGGTCTTCCTAATATGACTCAAGG	1080
307	I F K E A E K F F V S V G L P N M T Q G	326
1081	ATTCTGGGAAAATTCCATGCTAACGGACCCAGGAAATGTTTCAGAAAGCAGTCTGCCATCC	1140
327	F W E N S M L T D P G N V Q K A V C H P	346
1141	CACAGCTTGGGACCTGGGGAAGGGCGACTTCAGGATCCTTATGTGCACAAAGGTGACAAT	1200
347	T A W D L G K G D F R I L M C T K V T M	366
1201	GGACGACTTCCTGACAGCTCATCATGAGATGGGGCATATCCAGTATGATATGGCATATGC	1260
367	D D F L T A H H E M G H I Q Y D M A Y A	386
1261	TGCACAACCTTTTCTGCTAAGAAATGGAGCTAATGAAGGATTCCATGAAGCTGTTGGGGA	1320
387	A Q P F L L R N G A N E G F H E A V G E	406
1321	AATCATGTCACTTTTCTGCAGCCACACCTAAGCATTTAAATCCATTGGTCTTCTGTCACC	1380
407	I M S L S A A T P K H L K S I G L L S P	426
1381	CGATTTTCAAGAAGACAATGAAACAGAAATAAACTTCCTGCTCAAACAAGCACTCACGAT	1440
427	D F Q E D N E T E I N F L L K Q A L T I	446
1441	TGTTGGGACTCTGCCATTTACTTACATGTTAGAGAAGTGGAGGTGGATGGTCTTTAAAGG	1500
447	V G T L P F T Y M L E K W R W M V F K G	466
1501	GGAAATTCCCAAAGACCAGTGGATGAAAAAGTGGTGGGAGATGAAGCGAGAGATAGTTGG	1560
467	E I P K D Q W M K K W W E M K R E I V G	486
1561	GGTGGTGGAACCTGTGCCCCATGATGAAACATACTGTGACCCCGCATCTCTGTTCCATGT	1620
487	V V E P V P H D E T Y C D P A S L F H V	506
1621	TTCTAATGATTACTCATTCATTTCGATATTACACAAGGACCCTTTACCAATTCCAGTTTCA	1680
507	S N D Y S F I R Y Y T R T L Y Q F Q F Q	526
1681	AGAAGCACTTTGTCAAGCAGCTAAACATGAAGGCCCTCTGCACAAATGTGACATCTCAAA	1740
527	E A L C Q A A K H E G P L H K C D I S N	546

Figure 32C

```

1741 CTCTACAGAAGCTGGACAGAACTGTTCAATATGCTGAGGCTTGGAAAATCAGAACCCTG 1800
547 S T E A G Q K L F N M L R L G K S E P W 566

1801 GACCCTAGCATTGGAAAATGTTGTAGGAGCAAAGAACATGAATGTAAGGCCACTGCTCAA 1860
567 T L A L E N V V G A K N M N V R P L L N 586

1861 CTACTTTGAGCCCTTATTTACCTGGCTGAAAGACCAGAACAAGAATTCTTTTGTGGGATG 1920
587 Y F E P L F T W L K D Q N K N S F V G W 606

1921 GAGTACCGACTGGAGTCCATATGCAGACCAAAGCATCAAAGTGAGGATAAGCCTAAAAATC 1980
607 S T D W S P Y A D Q S I K V R I S L K S 626

1981 AGCTCTTGGAGATAAAGCATATGAATGGAACGACAATGAAATGTACCTGTTCCGATCATC 2040
627 A L G D K A Y E W N D N E M Y L F R S S 646

2041 TGTTCATATGCTATGAGGCAGTACTTTTAAAAAGTAAAAAATCAGATGATTCTTTTGG 2100
647 V A Y A M R Q Y F L K V K N Q M I L F G 666

2101 GGAGGAGGATGTGCGAGTGGCTAATTTGAAACCAAGAATCTCCTTTAATTTCTTTGTCAC 2160
667 E E D V R V A N L K P R I S F N F F V T 686

2161 TGCACCTAAAAATGTGTCTGATATCATTCCTAGAACTGAAGTTGAAAAGGCCATCAGGAT 2220
687 A P K N V S D I I P R T E V E K A I R M 706

2221 GTCCCGGAGCCGTATCAATGATGCTTTCCGTCTGAATGACAACAGCCTAGAGTTTCTGGG 2280
707 S R S R I N D A F R L N D N S L E F L G 726

2281 GATACAGCCAACACTTGGACCTCCTAACCAGCCCCCTGTTTCCATATGGCTGATTGTTTT 2340
727 I Q P T L G P P N Q P P V S I W L I V F 746

2341 TGGAGTTGTGATGGGAGTGATAGTGGTTGGCATTGTGCATCCTGATCTTCACTGGGATCAG 2400
747 G V V M G V I V V G I V I L I F T G I R 766

2401 AGATCGGAAGAAGAAAAATAAAGCAAGAAGTGAGAAAAATCCTTATGCCTCCATCGATAT 2460
767 D R K K K N K A R S G E N P Y A S I D I 786

2461 TAGCAAAGGAGAAAAATAATCCAGGATTCCAAAACACTGATGATGTTTCAGACCTCCTTTTA 2520
787 S K G E N N P G F Q N T D D V Q T S F * 806

2521 GAAAAATCTATGTTTTCTCTTGAGGTGATTTTGTGTATGTAAATGTTAATTTTCATGG 2580

2581 TATAGAAAAATATAAGATGATAAAGATATCATTAATGTCAAAACTATGACTCTGTTCAGA 2640

```

Figure 32D

2641 AAAAAAATTGTCCAAAGACAACATGGCCAAGGAGAGAGCATCTTCATTGACATTGCTTTC 2700
 2701 AGTATTTATTTCTGTCTCTGGATTGACTTCTGTTCTGTTTCTTAATAAGGATTTTGTAT 2760
 2761 TAGAGTATATTAGGGAAAGTGTGTATTTGGTCTCACAGGCTGTTTCAGGGATAATCTAAAT 2820
 2821 GTAAATGTCTGTTGAATTTCTGAAGTTGAAAACAAGGATATATCATTGGAGCAAGTGTG 2880
 2881 GATCTTGATGGAATATGGATGGATCACTTGTAAGGACAGTGCCTGGGAAGTGGTGTAGC 2940
 2941 TGCAAGGATTGAGAATGGCATGCATTAGCTCACTTTCATTTAATCCATTGTCAAGGATGA 3000
 3001 CATGCTTTCTTCACAGTAACTCAGTTCAAGTACTATGGTGATTGCTTACAGTGATGTTT 3060
 3061 GGAATCGATCATGCTTCTCTCAAGGTGACAGGTCTAAAGAGAGAAGAATCCAGGGAACAG 3120
 3121 GTAGAGGACATTGCTTTTTCACTTCCAAGGTGCTTGATCAACATCTCCCTGACAACACAA 3180
 3181 AACTAGAGCCAGGGGCCTCCGTGAACTCCCAGAGCATGCCTGATAGAACTCATTTCTAC 3240
 3241 TGTTCCTAACTGTGGAGTGAATGGAAATTCCTCAACTGTATGTTACCCCTCTGAAGTGGGT 3300
 3301 ACCCAGTCTCTTAAATCTTTTGTATTTGCTCACAGTGTTTGAGCAGTGCTGAGCACAAAG 3360
 3361 CAGACACTCAATAAATGCTAGATTTACACACTCAAAAAAAAAAAAA 3405

Figure 33A

1	ATGCATCTTATCGACTACCTGCTCCTCCTGCTGGTTGGACTACTGGCCCTTTCTCATGGC	60
1	M H L I D Y L L L L V G L L A L S H G	20
61	CAGCTGCACGTTGAGCATGATGGTGAGAGTTGCAGTAACAGCTCCCACCAGCAGATTCTG	120
21	Q L H V E H D G E S C S N S S H Q Q I L	40
121	GAGACAGGTGAGGGCTCCCCAGCCTCAAGATAGCCCCCTGCCAATGCTGACTTTGCCTTC	180
41	E T G E G S P S L K I A P A N A D F A F	60
181	CGCTTCTACTACCTGATCGCTTCGGAGACCCCCGGGGAAGAACATCTTTTCTCCCCGCTG	240
61	R F Y Y L I A S E T P G K N I F F S P L	80
241	AGCATCTCGGCGGCCTACGCCATGCTTTCCCTGGGGGCCTGCTCACACAGCCGCAGCCAG	300
81	S I S A A Y A M L S L G A C S H S R S Q	100
301	ATCCTTGAGGGCCTGGGCTTCAACCTCACCGAGCTGTCTGAGTCCGATGTCCATAGGGGC	360
101	I L E G L G F N L T E L S E S D V H R G	120
361	TTCCAGCACCTCCTGCACACTCTCAACCTCCCCGGCCATGGGCTGGAAACACGCGTGGGC	420
121	F Q H L L H T L N L P G H G L E T R V G	140
421	AGTGCCTCTGTTCTGAGCCACAACCTGAAGTTCCTTGCAAAATTCTGAATGACACCATG	480
141	S A L F L S H N L K F L A K F L N D T M	160
481	GCCGTCCTATGAGGCTAAACTCTTCCACACCAACTTCTACGACACTGTGGGCACAATCCAG	540
161	A V Y E A K L F H T N F Y D T V G T I Q	180
541	CTTATCAACGACCACGTCAAGAAGGAAACTCGAGGGAAGATTGTGGATTGGTTCAGTGAG	600
181	L I N D H V K K E T R G K I V D L V S E	200
601	CTCAAGAAGGACGTCTTGATGGTGCTGGTGAATTACATTTACTTCAAAGCCCTGTGGGAG	660
201	L K K D V L M V L V N Y I Y F K A L W E	220
661	AAACCATTCAATTCCTCAAGGACCACTCCCAAAGACTTCTATGTTGATGAGAACACAACA	720
221	K P F I S S R T T P K D F Y V D E N T T	240
721	GTCCGGGTGCCCATGATGCTGCAGGACCAGGAGCATCACTGGTATCTTCATGACAGATAC	780
241	V R V P M M L Q D Q E H H W Y L H D R Y	260
781	TTGCCCTGCTCGGTGCTACGGATGGATTACAAAGGAGACGCAACCGTGTTCATTCTC	840
261	L P C S V L R M D Y K G D A T V F F I L	280

Figure 33B

841 CCTAACCAAGGCCAAAATGAGGGAGATTGAAGAGGTTCTGACTCCAGAGATGCTAATGAGG 900
281 P N Q G K M R E I E E V L T P E M L M R 300

901 TGGAACAACCTTGTTGCGGAAGAGGAATTTTACAAGAAGCTAGAGTTGCATCTTCCCAAG 960
301 W N N L L R K R N F Y K K L E L H L P K 320

961 TTCTCCATTTCTGGCTCCTATGTATTAGATCAGATTTTGCCCAGGCTGGGCTTCACGGAT 1020
321 F S I S G S Y V L D Q I L P R L G F T D 340

1021 CTGTTCTCCAAGTGGGCTGACTTATCCGGCATCACCAAACAGCAAAAACTGGAGGCATCC 1080
341 L F S K W A D L S G I T K Q Q K L E A S 360

1081 AAAAGTTTCCACAAGGCCACCTTGGACGTGGATGAGGCTGGCACCAGAGGCTGCAGCAGCC 1140
361 K S F H K A T L D V D E A G T E A A A A 380

1141 ACCACGTTTCGCGATCAAATTCTTCTCTGCCCAGACCAATCGCCACATCCTGCGATTCAAC 1200
381 T T F A I K F F S A Q T N R H I L R F N 400

1201 CGGCCCTTCCTTGTGGTGATCTTTTCCACCAGCACCAGAGTGTCTCTTTCTGGGCAAG 1260
401 R P F L V V I F S T S T Q S V L F L G K 420

1261 GTCGTCGACCCACGAAACCATAG 1284
421 V V D P T K P * 428

Figure 34A

1	ATGCATCTTATCGACTACCTGCTCCTCCTGCTGGTTGGACTACTGGCCCTTTCTCATGGC	60
1	M H L I D Y L L L L L V G L L A L S H G	20
61	CAGCTGCACGTTGAGCATGATGGTGAGAGTTGCAGTAACAGCTCCCACCAGCAGATTCTG	120
21	Q L H V E H D G E S C S N S S H Q Q I L	40
121	GAGACAGGTGAGGGCTCCCCAGCCTCAAGATAGCCCCTGCCAATGCTGACTTTGCCTTC	180
41	E T G E G S P S L K I A P A N A D F A F	60
181	CGCTTCTACTACCTGATCGCTTCGGAGACCCCGGGGAAGAACATCTTTTTCTCCCCGCTG	240
61	R F Y Y L I A S E T P G K N I F F S P L	80
241	AGCATCTCGGCGGCCCTACGCCATGCTTTCCCTGGGGGCCTGCTCACACAGCCGCAGCCAG	300
81	S I S A A Y A M L S L G A C S H S R S Q	100
301	ATCCTTGAGGGCCTGGGCTTCAACCTCACCGAGCTGTCTGAGTCCGATGTCCATAGGGGC	360
101	I L E G L G F N L T E L S E S D V H R G	120
361	TTCCAGCACCTCCTGCACACTCTCAACCTCCCCGGCCATGGGCTGGAAACACGCGTGGGC	420
121	F Q H L L H T L N L P G H G L E T R V G	140
421	AGTGCTCTGTTCCTGAGCCACAACCTGAAGTTCCTTGCAAAATTCCTGAATGACACCATG	480
141	S A L F L S H N L K F L A K F L N D T M	160
481	GCCGTCTATGAGGCTAAACTCTTCCACACCAACTTCTACGACACTGTGGGCACAATCCAG	540
161	A V Y E A K L F H T N F Y D T V G T I Q	180
541	CTTATCAACGACCACGTCAAGAAGGAAACTCGAGGGAAGATTGTGGATTGGTCAGTGAG	600
181	L I N D H V K K E T R G K I V D L V S E	200
601	CTCAAGAAGGACGTCTTGATGGTGCTGGTGAATTACATTTACTTCAAAGCCCTGTGGGAG	660
201	L K K D V L M V L V N Y I Y F K A L W E	220
661	AAACCATTTCATTTCTCAAGGACCACTCCCAAAGACTTTTATGTTGATGAGAACACAACA	720
221	K P F I S S R T T P K D F Y V D E N T T	240
721	GTCCGGGTGCCCATGATGCTGCAGGACCAGGAGCATCACTGGTATCTTCATGACAGATAC	780
241	V R V P M M L Q D Q E H H W Y L H D R Y	260
781	TTGCCCTGCTCGGTGCTACGGATGGATTACAAAGGAGACGCAACCGTGTTTTTCATTCTC	840
261	L P C S V L R M D Y K G D A T V F F I L	280

Figure 34B

841 CCTAACCAAGGCAAAATGAGGGAGATTGAAGAGGTTCTGACTCCAGAGATGCTAATGAGG 900
281 P N Q G K M R E I E E V L T P E M L M R 300

901 TGGAACAACCTTGTTCGGAAGAGGAATTTTACAAGAAGCTAGAGTTGCATCTTCCCAAG 960
301 W N N L L R K R N F Y K K L E L H L P K 320

961 TTCTCCATTTCTGGCTCCTATGTATTAGATCAGATTTTGCCCAGGCTGGGCTTCACGGAT 1020
321 F S I S G S Y V L D Q I L P R L G F T D 340

1021 CTGTTCTCCAAGTGGGCTGACTTATCCGGCATCACCAAACAGCAAAAACCTGGAGGCATCC 1080
341 L F S K W A D L S G I T K Q Q K L E A S 360

1081 AAAAGTTTCCACAAGGCCACCTTGGACGTGGATGAGGCTGGCACCAGAGGCTGCAGCAGCC 1140
361 K S F H K A T L D V D E A G T E A A A A 380

1141 ACCACGTTCGCGATCAAATTCTTCTCTGCCCAGACCAATCGCCACATCCTGCGATTCAAC 1200
381 T T F A I K F F S A Q T N R H I L R F N 400

1201 CGGCCCTTCCTTGTGGTGATCTTTTCCACCAGCACCCAGAGTGTCTCTTTCTGGGCAAG 1260
401 R P F L V V I F S T S T Q S V L F L G K 420

1261 GTCGTCGACCCACGAAACCATAG 1284
421 V V D P T K P * 428

Figure 35A

1	ATGCATCTTATCGACTACCTGCTCCTCCTGCTGGTTGGACTACTGGCCCTTTCTCATGGC	60
1	M H L I D Y L L L L L V G L L A L S H G	20
61	CAGCTGCACGTTGAGCATGATGGTGAGAGTTGCAGTAACAGCTCCCACCAGCAGATTCTG	120
21	Q L H V E H D G E S C S N S S H Q Q I L	40
121	GAGACAGGTGAGGGCTCCCCAGCCTCAAGATAGCCCCTGCCAATGCTGACTTTGCCTTC	180
41	E T G E G S P S L K I A P A N A D F A F	60
181	CGCTTCTACTACCTGATCGCTTCGGAGACCCCGGGGAAGAACATCTTTTTCTCCCCGCTG	240
61	R F Y Y L I A S E T P G K N I F F S P L	80
241	AGCATCTCGGCGGCCTACGCCATGCTTTCCTGGGGGCCTGCTCACACAGCCGCAGCCAG	300
81	S I S A A Y A M L S L G A C S H S R S Q	100
301	ATCCTTGAGGGCCTGGGCTTCAACCTCACCGAGCTGTCTGAGTCCGATGTCCATAGGGGC	360
101	I L E G L G F N L T E L S E S D V H R G	120
361	TTCCAGCACCTCCTGCACACTCTCAACCTCCCCGGCCATGGGCTGGAAACACGCGTGGGC	420
121	F Q H L L H T L N L P G H G L E T R V G	140
421	AGTGCTCTGTTCTGAGCCACAACCTGAAGTTCCTTGCAAAATTCCTGAATGACACCATG	480
141	S A L F L S H N L K F L A K F L N D T M	160
481	GCCGTCATAGAGGCTAAACTCTTCCACACCAACTTCTACGACACTGTGGGCACAATCCAG	540
161	A V Y E A K L F H T N F Y D T V G T I Q	180
541	CTTATCAACGACCACGTCAAGAAGGAAACTCGAGGGAAGATTGTGGATTGGTCAGCGAG	600
181	L I N D H V K K E T R G K I V D L V S E	200
601	CTCAAGAAGGACGTCTTGATGGTGCTGGTGAATTACATTTACTTCAAAGCCCTGTGGGAG	660
201	L K K D V L M V L V N Y I Y F K A L W E	220
661	AAACCATTCAATTCCTCAAGGACCACTCCCAAAGACTTCTATGTTGATGAGAACACAACA	720
221	K P F I S S R T T P K D F Y V D E N T T	240
721	GTCCGGGTGCCCATGATGCTGCAGGACCAGGAGCATCACTGGTATCTTCATGACAGATAC	780
241	V R V P M M L Q D Q E H H W Y L H D R Y	260
781	TTGCCCTGCTCGGTGCTACGGATGGATTACAAAGGAGACGCAACCGTGTTTTTCATTCTC	840
261	L P C S V L R M D Y K G D A T V F F I L	280

Figure 35B

841 CCTAACCAAGGCAAAATGAGGGAGATTGAAGAGGTTCTGACTCCAGAGATGCTAATGAGG 900
281 P N Q G K M R E I E E V L T P E M L M R 300

901 TGGAACAACTTGTTCGCGAAGAGGAATTTTACAAGAAGCTAGAGTTGCATCTTCCAAG 960
301 W N N L L R K R N F Y K K L E L H L P K 320

961 TTCTCCATTTCTGGCTCCTATGTATTAGATCAGATTTTGCCAGGCTGGGCTTCACGGAT 1020
321 F S I S G S Y V L D Q I L P R L G F T D 340

1021 CTGTTCTCCAAGTGGGCTGACTTATCCGGCATCACCAAACAGCAAAAAGTGGAGGCATCC 1080
341 L F S K W A D L S G I T K Q Q K L E A S 360

1081 AAAAGTTTCCACAAGGCCACCTTGGACGTGGATGAGGCTGGCACCGAGGCTGCAGCAGCC 1140
361 K S F H K A T L D V D E A G T E A A A A 380

1141 ACCACGTTTCGCGATCAAATTCTTCTCTGCCCAGACCAATCGCCACATCCTGCGATTCAAC 1200
381 T T F A I K F F S A Q T N R H I L R F N 400

1201 CGGCCCTTCCTTGTGGTGATCTTTTCCACCAGCACCCAGAGTGTCTCTTTCTGGGCAAG 1260
401 R P F L V V I F S T S T Q S V L F L G K 420

1261 GTCGTCGACCCACGAAACCATAG 1284
421 V V D P T K P * 428

Figure 36A

1	ATGCATCTTATCGACTACCTGCTCCTCCTGCTGGTTGGACTACTGGCCCTTTCTCATGGC	60
1	M H L I D Y L L L L L V G L L A L S H G	20
61	CAGCTGCACGTTGAGCATGATGGTGAGAGTTGCAGTAACAGCTCCCACCAGCAGATTCTG	120
21	Q L H V E H D G E S C S N S S H Q Q I L	40
121	GAGACAGGTGAGGGCTCCCCAGCCTCAAGATAGCCCCTGCCAATGCTGACTTTGCCTTC	180
41	E T G E G S P S L K I A P A N A D F A F	60
181	CGCTTCTACTACCTGATCGCTTCGGAGACCCCCGGGGAAGAACATCTTTTTCTCCCCGCTG	240
61	R F Y Y L I A S E T P G K N I F F S P L	80
241	AGCATCTCGGCGGCCTACGCCATGCTTTCCCTGGGGGCCTGCTCACACAGCCGCAGCCAG	300
81	S I S A A Y A M L S L G A C S H S R S Q	100
301	ATCCTTGAGGGCCTGGGCTTCAACCTCACCGAGCTGTCTGAGTCCGATGTCCATAGGGGC	360
101	I L E G L G F N L T E L S E S D V H R G	120
361	TTCCAGCACCTCCTGCACACTCTCAACCTCCCCGCCATGGGCTGGAAACACGCGTGGGC	420
121	F Q H L L H T L N L P G H G L E T R V G	140
421	AGTGCTCTGTTCCCTGAGCCACAACCTGAAGTTCCTTGCAAAATTCCTGAATGACACCATG	480
141	S A L F L S H N L K F L A K F L N D T M	160
481	GCCGTCTATGAGGCTAAACTCTTCCACACCAACTTCTACGACACTGTGGGCACAATCCAG	540
161	A V Y E A K L F H T N F Y D T V G T I Q	180
541	CTTATCAACGACCACGTCAAGAAGGAAACTCGAGGGAAGATTGTGGATTGGTTCAGNGAG	600
181	L I N D H V K K E T R G K I V D L V S E	200
601	CTCAAGAAGGACGTCTTGATGGTGCTGGTGAATTACATTTACTTCAAAGCCCTGTGGGAG	660
201	L K K D V L M V L V N Y I Y F K A L W E	220
661	AAACCATTCAATTCCTCAAGGACCACTCCCAAAGACTTNTATGTTGATGAGAACACAACA	720
221	K P F I S S R T T P K D F Y V D E N T T	240
721	GTCCGGGTGCCCATGATGCTGCAGGACCAGGAGCATCACTGGTATCTTCATGACAGATAC	780
241	V R V P M M L Q D Q E H H W Y L H D R Y	260
781	TTGCCCTGCTCGGTGCTACGGATGGATTACAAAGGAGACGCAACCGTGTTTTTCATTCTC	840
261	L P C S V L R M D Y K G D A T V F F I L	280

Figure 36B

841	CCTAACCAAGGCAAAATGAGGGAGATTGAAGAGGTTCTGACTCCAGAGATGCTAATGAGG	900
281	P N Q G K M R E I E E V L T P E M L M R	300
901	TGGAACAACTTGTTGCGGAAGAGGAATTTTACAAGAAGCTAGAGTTGCATCTTCCCAAG	960
301	W N N L L R K R N F Y K K L E L H L P K	320
961	TTCTCCATTTCTGGCTCCTATGTATTAGATCAGATTTTGCCCAGGCTGGGCTTCACGGAT	1020
321	F S I S G S Y V L D Q I L P R L G F T D	340
1021	CTGTTCTCCAAGTGGGCTGACTTATCCGGCATCACCAAACAGCAAAAACCTGGAGGCATCC	1080
341	L F S K W A D L S G I T K Q Q K L E A S	360
1081	AAAAGTTTCCACAAGGCCACCTTGGACGTGGATGAGGCTGGCACCGAGGCTGCAGCAGCC	1140
361	K S F H K A T L D V D E A G T E A A A A	380
1141	ACCACGTTTCGCGATCAAATTCTTCTCTGCCCAGACCAATCGCCACATCCTGCGATTCAAC	1200
381	T T F A I K F F S A Q T N R H I L R F N	400
1201	CGGCCCTTCCTTGTGGTGATCTTTTCCACCAGCACCCAGAGTGTCTCTTTCTGGGCAAG	1260
401	R P F L V V I F S T S T Q S V L F L G K	420
1261	GTCGTCGACCCACGAAACCATAG	1284
421	V V D P T K P *	428

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 37A

1	CGCCCAACCCAAGTTCAAAGGCTGATAAGAGAGAAAATCTCATGAGGAGGTTTTAGTCTA	60
61	GGGAAAGTCATTTCAGTGGATGTGATCTTGGCTCACAGGGGACGATGTCAAGCTCTTCCTG	120
1	M S S S S W	6
121	GCTCCTTCTCAGCCTTGTTGCTGTAAGTCTGCTCAGTCCACCATTGAGGAACAGGCCAA	180
7	L L L S L V A V T A A Q S T I E E Q A K	26
181	GACATTTTGGACAAGTTTAACCACGAAGCCGAAGACCTGTTCTATCAAAGTTCACTTGC	240
27	T F L D K F N H E A E D L F Y Q S S L A	46
241	TTCTTGGAATTATAACACCAATATTACTGAAGAGAATGTCCAAAACATGAATAATGCTGG	300
47	S W N Y N T N I T E E N V Q N M N N A G	66
301	GGACAAATGGTCTGCCTTTTAAAGGAACAGTCCACACTTGCCCAAATGTATCCACTACA	360
67	D K W S A F L K E Q S T L A Q M Y P L Q	86
361	AGAAATTCAGAATCTCACAGTCAAGCTTCAGCTGCAGGCTCTTCAGCAAAATGGGTCTTC	420
87	E I Q N L T V K L Q L Q A L Q Q N G S S	106
421	AGTGCTCTCAGAAGACAAGAGCAAACGGTTGAACACAATTCTAAATACAATGAGCACCAT	480
107	V L S E D K S K R L N T I L N T M S T I	126
481	CTACAGTACTGGAAAAGTTTGTAACCCAGATAATCCACAAGAATGCTTATTACTTGAACC	540
127	Y S T G K V C N P D N P Q E C L L L E P	146
541	AGGTTTGAATGAAATAATGGCAAACAGTTTAGACTACAATGAGAGGCTCTGGGCTTGGGA	600
147	G L N E I M A N S L D Y N E R L W A W E	166
601	AAGCTGGAGATCTGAGGTCGGCAAGCAGCTGAGGCCATTATATGAAGAGTATGTGGTCTT	660
167	S W R S E V G K Q L R P L Y E E Y V V L	186
661	GAAAAATGAGATGGCAAGAGCAAATCATTATGAGGACTATGGGGATTATTGGAGAGGAGA	720
187	K N E M A R A N H Y E D Y G D Y W R G D	206
721	CTATGAAGTAAATGGGGTAGATGGCTATGACTACAGCCGCGGCCAGTTGATTGAAGATGT	780
207	Y E V N G V D G Y D Y S R G Q L I E D V	226
781	GGAACATACCTTTGAAGAGATTAAACCATTATATGAACATCTTCATGCCTATGTGAGGGC	840
227	E H T F E E I K P L Y E H L H A Y V R A	246
841	AAAGTTGATGAATGCCTATCCTTCCTATATCAGTCCAATTGGATGCCTCCCTGCTCATTT	900
247	K L M N A Y P S Y I S P I G C L P A H L	266

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 37B

901	GCTTGGTGATATGTGGGGTAGATTTTGGACAAATCTGTACTCTTTGACAGTTCCCTTTGG	960
267	L G D M W G R F W T N L Y S L T V P F G	286
961	ACAGAAACCAAACATAGATGTTACTGATGCAATGGTGGACCAGGCCTGGGATGCACAGAG	1020
287	Q K P N I D V T D A M V D Q A W D A Q R	306
1021	AATATTCAAGGAGGCCGAGAAGTTCTTTGTATCTGTTGGTCTTCCTAATATGACTCAAGG	1080
307	I F K E A E K F F V S V G L P N M T Q G	326
1081	ATTCTGGGAAAATTCCATGCTAACGGACCCAGGAAATGTTTCAGAAAGCAGTCTGCCATCC	1140
327	F W E N S M L T D P G N V Q K A V C H P	346
1141	CACAGCTTGGGACCTGGGAAGGGCGACTTCAGGATCCTTATGTGCACAAAGGTGACAAT	1200
347	T A W D L G K G D F R I L M C T K V T M	366
1201	GGACGACTTCCTGACAGCTCATCATGAGATGGGGCATATCCAGTATGATATGGCATATGC	1260
367	D D F L T A H H E M G H I Q Y D M A Y A	386
1261	TGCACAACCTTTTCTGCTAAGAAATGGAGCTAATGAAGGATTCCATGAAGCTGTTGGGGA	1320
387	A Q P F L L R N G A N E G F H E A V G E	406
1321	AATCATGTCACTTTCTGCAGCCACACCTAAGCATTATAAATCCATTGGTCTTCTGTCACC	1380
407	I M S L S A A T P K H L K S I G L L S P	426
1381	CGATTTTCAAGAAGACAATGAAACAGAAATAAACTTCCTGCTCAAACAAGCACTCACGAT	1440
427	D F Q E D N E T E I N F L L K Q A L T I	446
1441	TGTTGGGACTCTGCCATTTACTTACATGTTAGAGAAGTGGAGGTGGATGGTCTTTAAAGG	1500
447	V G T L P F T Y M L E K W R W M V F K G	466
1501	GGAAATTCCCAAAGACCAGTGGATGAAAAAGTGGTGGGAGATGAAGCGAGAGATAGTTGG	1560
467	E I P K D Q W M K K W W E M K R E I V G	486
1561	GGTGGTGGAAACCTGTGCCCCATGATGAAACATACTGTGACCCCGCATCTCTGTTCCATGT	1620
487	V V E P V P H D E T Y C D P A S L F H V	506
1621	TTCTAATGATTACTCATTCATTCGATATTACACAAGGACCCTTTACCAATTCCAGTTTCA	1680
507	S N D Y S F I R Y Y T R T L Y Q F Q F Q	526
1681	AGAAGCACTTTGTCAAGCAGCTAAACATGAAGGCCCTCTGCACAAATGTGACATCTCAAA	1740
527	E A L C Q A A K H E G P L H K C D I S N	546

[illegible][illegible]

Figure 37D

2641 AAAAAAAAAATTGTCCAAAGACAACATGGCCAAGGAGAGAGCATCTTCATTGACATTGCTTTC 2700

2701 AGTATTTATTTCTGTCTCTGGATTTGACTTCTGTTCTGTTTCTTAATAAGGATTTTGTAT 2760

2761 TAGAGTATATTAGGGAAAGTGTGTATTTGGTCTCACAGGCTGTTTCAGGGATAATCTAAAT 2820

2821 GTAAATGTCTGTTGAATTTCTGAAGTTGAAAACAAGGATATATCATTGGAGCAAGTGTTG 2880

2881 GATCTTGTATGGAATATGGATGGATCACTTGTAAGGACAGTGCCTGGGAAGTGGTGTAGC 2940

2941 TGCAAGGATTGAGAATGGCATGCATTAGCTCACTTTCATTTAATCCATTGTCAAGGATGA 3000

3001 CATGCTTTCCTTCACAGTAACTCAGTTCAAGTACTATGGTGATTTGCCTACAGTGATGTTT 3060

3061 GGAATCGATCATGCTTTCCTCAAGGTGACAGGTCTAAAGAGAGAAGAATCCAGGGAACAG 3120

3121 GTAGAGGACATTGCTTTTTCACTTCCAAGGTGCTTGATCAACATCTCCCTGACAACACAA 3180

3181 AACTAGAGCCAGGGGCCTCCGTGAACTCCCAGAGCATGCCTGATAGAACTCATTTCTAC 3240

3241 TGTTCCTCTAACTGTGGAGTGAATGGAAATTCCAAGTGTATGTTACCCCTCTGAAGTGGGT 3300

3301 ACCCAGTCTCTTAAATCTTTGTATTTGCTCACAGTGTTTGAGCAGTGCTGAGCACAAAG 3360

3361 CAGACACTCAATAAATGCTAGATTTACACACTCAAAAAAAAAAAAA 3405

Figure 38A

```

1  ATGTTCTCTC CCTGGAAGAT ATCAATGTTT CTGTCTGTTN GTGAGGNCTC CGTGCCCACC 60
1  M F S P   W K I   S M F   L S V X   E X S   V P T   20

61  ACGGCCTCTT TCAGCGCCGA CATGTCTCAAT GTCACCTTGC AAGGGCCCAC TCTTAACGGG 120
21  T A S F   S A D   M L N   V T L Q   G P T   L N G   40

121 ACCTTTGCCC AGAGCAAATG CCCCCAAGTG GAGTGGCTGG GCTGGCTCAA CACCATCCAG 180
41  T F A Q   S K C   P Q V   E W L G   W L N   T I Q   60

181 CCCCCCTTCC TCTGGGTGCT GTTCGTGCTG GCCACCCTAG AGAACATCTT TGTCCTCAGC 240
61  P P F L   W V L   F V L   A T L E   N I F   V L S   80

241 GTCTTCTGCC TGCACAAGAG CAGCTGCACG GTGGCAGAGA TCTACCTGGG GAACCTGGCC 300
81  V F C L   H K S   S C T   V A E I   Y L G   N L A   100

301 GCAGCAGACC TGATCCTGGC CTGCGGGCTG CCCTTCTGGG CCATCACCAT CTCCAACAAC 360
101 A A D L   I L A   C G L   P F W A   I T I   S N N   120

361 TTCGACTGGC TCTTTGGGGA GACGCTCTGC CGCGTGGTGA ATGCCATTAT CTCCATGAAC 420
121 F D W L   F G E   T L C   R V V N   A I I   S M N   140

421 CTGTACAGCA GCATCTGTTT CCTGATGCTG GTGAGCATCG ACCGCTACCT GGCCCTGGTG 480
141 L Y S S   I C F   L M L   V S I D   R Y L   A L V   160

481 AAAACCATGT CCATGGGCCG GATGCGCGGC GTGCGCTGGG CCAAGCTCTA CAGCTTGGTG 540
161 K T M S   M G R   M R G   V R W A   K L Y   S L V   180

541 ATCTGGGGGT GTACGCTGCT CCTGAGCTCA CCCATGCTGG TGTCCGGAC CATGAAGGAG 600
181 I W G C   T L L   L S S   P M L V   F R T   M K E   200

601 TACAGCGATG AGGGCCACAA CGTCACCGCT TGTGTCATCA GCTACCCATC CCTCATCTGG 660
201 Y S D E   G H N   V T A   C V I S   Y P S   L I W   220

661 GAAGTGTTCA CCAACATGCT CCTGAATGTC GTGGGCTTCC TGCTGCCCCT GAGTGTATC 720
221 E V F T   N M L   L N V   V G F L   L P L   S V I   240

721 ACCTTCTGCA CGATGCAGAT CATGCAGGTG CTGCGGAACA ACGAGATGCA GAAGTTCAAG 780
241 T F C T   M Q I   M Q V   L R N N   E M Q   K F K   260

781 GAGATCCAGA CGGAGAGGAG GGCCACGGTG CTAGTCCTGG TTGTGCTGCT GCTATTCATC 840
261 E I Q T   E R R   A T V   L V L V   V L L   L F I   280

841 ATCTGCTGGC TGCCCTTCCA GATCAGCACC TTCCTGGATA CGCTGCATCG CCTCGGCATC 900
281 I C W L   P F Q   I S T   F L D T   L H R   L G I   300

```

Figure 38B

```

901 CTCTCCAGCT GCCAGGACGA GCGCATCATC GANGTAATCA CACAGATCGC CTCCTTCATG 960
301 L S S C Q D E R I I D V I T Q I A S F M 320

961 GCCTACAGCA ACAGCTGCCT CAACCCACTG GTGTACGTGA TCGTGCGCAA GCGCTTCCGA 1020
321 A Y S N S C L N P L V Y V I V G K R F R 340

1021 AAGAAGTCTT GGGAGGTGTA CCAGGGAGTG TGCCAGAAAG NGGGCTGCAG GTCAGAACCC 1080
341 K K S W E V Y Q G V C Q K X G C R S E P 360

1081 ATTCAGATGG AGAACTCCAT GGGCACACTG CGGACCTCCA TCTCCGTGGA ACGCCAGATT 1140
361 I Q M E N S M G T L R T S I S V E R Q I 380

1141 CACAACTGC AGGACTGGGC AGGGAGCAGA CAGTGAGCAA ACGCCAGCAG GGCTGCTGTG 1200
381 H K L Q D W A G S R Q * 391

1201 AATTTGTGTA AGGATTGAGG GACAGTTGCT TTTCAGCATG GGCCCAGGAA TGCCAAGGAG 1260

1261 ACATCTATGC ACGACCTTGG GAAATGAGTT GATGTCTCCG GTAAACACC GGAGACTAAT 1320

1321 TCCTGCCCTG CCAATTTTG CAGGGAGCAT GGCTGTGAGG ATGGGGTGAA CTCACGCACA 1380

1381 GCCAAGGACT CCAAATCAC AACAGCATTA CTGTTCTTAT TTGCTGCCAC ACCTGAGCCA 1440

1441 GCCTGCTCCT TCCCAGGAGT GGAGGAGGCC TGGGGGAGG GAGAGGAGTG ACTGAGCTTC 1500

1501 CCTCCCCTGT GTTCTCCGTC CCTGCCCCAG CAAGACAAC TAGATCTCCA GGAGAACTGC 1560

1561 CATCCAGCTT TGGTGCAATG GCTGAGTGCA CAAGTGAGTT GTTGCCCTGG GTTTCTTTAA 1620

1621 TCTATTCAGC TAGAACTTTG AAGGACAATT TCTTGCAATTA ATAAAGGTTA AGCCCTGAGG 1680

1681 GGTCCCTGAT AACAACTGG AGACCAGGAT TTTATGGCTC CCCTCACTGA TGGACAAGGA 1740

1741 GGTCTGTGCC AAAGAAGAAT CCAATAAGCA CATATTGAGC ACTTGCTGTA TATGCAGTAT 1800

1801 TGAGCACTGT AGGCAAGACC CAAGAAAGAG AAGGAGCCAT CTCCATCTTG AAGGAACTCA 1860

1861 AAGACTCAAG TGGGAACGAC TGGGCACTGC CACCACCAGA AAGCTGTTCTG ACGAGACGGT 1920

1921 CGAGCAGGGT GCTGTGGGTG ATATGGACAG CAGAAGGGGG AGACCAAGGT TCCAGCTCAA 1980

1981 CCAATAACTA TTGCACAACC ACCTGTCCCT GCCTCAGTTC CCTTTTATGT AACATGAAGT 2040

```

Figure 38C

2041 CGTTGTGAGG GTTAAAGGCA GTAACAGGTA TAAAGTACTT AGAAAAGCAA AGGGTGCTAC 2100
 2101 GTACATGTGA GGCATCATTÀ CGCAGACGTA ACTGGGATAT GTTTACTATA AGGAAAAGAC 2160
 2161 ACTGAGGTCT AGAAATAGCT CCGTGGAGCA GAATCAGTAT TGGGAGCCGG TGGCGGTGTG 2220
 2221 AAGCACCAGT GTCTGGCACA CAGTAGGTGC TCATTGGCTC CCTTCCACCT GTCATTCCCA 2280
 2281 CCACCCTGAG GCCCCAACCG CCACACACAC AGGAGCATT TGGAGAGAAG CCATGTCTTC 2340
 2341 AAAGTCTGAT TTGTGATGAG GCAGAGGAAG ATATTTCTAA TCGGTCTTGC CCAGAGGATC 2400
 2401 ACAGTGCTGA GACCCCCAC CACCAGCCGG TACCTGGGAA GGGGGAGAGT GCAGGCCTGC 2460
 2461 TCAGGGACTG TTCCTGTCTC AGCAACCAAG GGATTGTTCC TGTCAATCAA TGGTTTATTG 2520
 2521 GAAGGTGGCC CAGTATGAGC CCTAGAAGAG TGTGAAAAGG AATGGCAATG GTGTTACCA 2580
 2581 TCGGCAGTGC CAGGGCAGCA CTCATTCACT TGATAAATGA ATATTTATTA GCTGGTTGGA 2640
 2641 GAGCTAGAAC CTGGAGAGCT AGAACCTGGA GAACTAGAAC CTGGAGGGCT AGAACCTGGA 2700
 2701 GAGGCTAGAA CCAAGAAGGG CTAGAACCTG GAGGGGCTAG AACCTAGAGA AGCTAAAACC 2760
 2761 TGAGCTAGAA GCTGGAGGAC TAGAACCTGG AGGGCTGGAA TCTGAAGGGC TAGAACCTGG 2820
 2821 AGGGCTGGAA TCTGGAGAGC TAGAACCTGG AGGGCTAGAA CCTGGAGGGC TAGAACCTAG 2880
 2881 AAGGGCTAGA ACCTGGAGGG CTGGAATCTG GAGAGCTAGA ACCTGGAGGG CTAGAACCTG 2940
 2941 GAGGGCTAGA ACCTAGAAGG GCTAGAACCT GGAGGGCTAG AACCTGGCAG GTTAGAACCT 3000
 3001 AGAAGGGCTA GAACCTGGAG AGCCAGAACC TGGAGGGCTA GAACCTGGAA GGGCTAGAAC 3060
 3061 CTGTAGAGCT AGAACATGGA GAGCTAGAAC CCGGCAGGCT AGAACCTGGC AAGCTAGAAC 3120
 3121 CTGGAGGGAA TGAACCTGGA GGGCTAGAAC CTGGAGAATG AGAAAAATTT ACATGGCAAÀ 3180
 3181 GAGCCCATAA ATCCTGACCA ATCCAACCTCT GAATTTTAAA GCAAAAAGCGT GAAAAAAAAG 3240

Figure 38D

3241 ATTCCCTCCT TACCCCAAC CCACTCTTTT TCCCACCAC CCACTCTCCT CTGCCTCAGT 3300
3301 AAGTATCTGG AGGAAGAAA CAGGTGAAAG AAGAAGTAA AACCATTAG TATTAGTATT 3360
3361 AGAATGAAGT CAAACTGTGC CACACATGGT GAATGAAAA AAAAAAAG AGGCTGTGTT 3420
3421 TTGTCACACA GGCAGTCAT TCAGCACCAG AGCACGTGAT GGTCTGAGAC TCTCTTAGGA 3480
3481 GCAGAGCTCT GCCGCAATGG CCATGTGGGG ATCCACACCT GGTCTGAGGG GCAACTGAGT 3540
3541 CTGCGGGAGA AGAGCGGCC TATGCATGGT GTAGATGCCC TGATAAAGAA CATCTGTCCT 3600
3601 GTGAAAGACT CAATGAGCTG TTATGTTGTA AACAGGAAGC ATTTACATC CAAACGAGAA 3660
3661 AATCATGTAA ACATGTGTCT TTTCTGTAGA GCATAATAAA TGGATGAGGT TTTTGCAAAA 3720
3721 AAAAAAAAA AAA 3733

D0053 NP

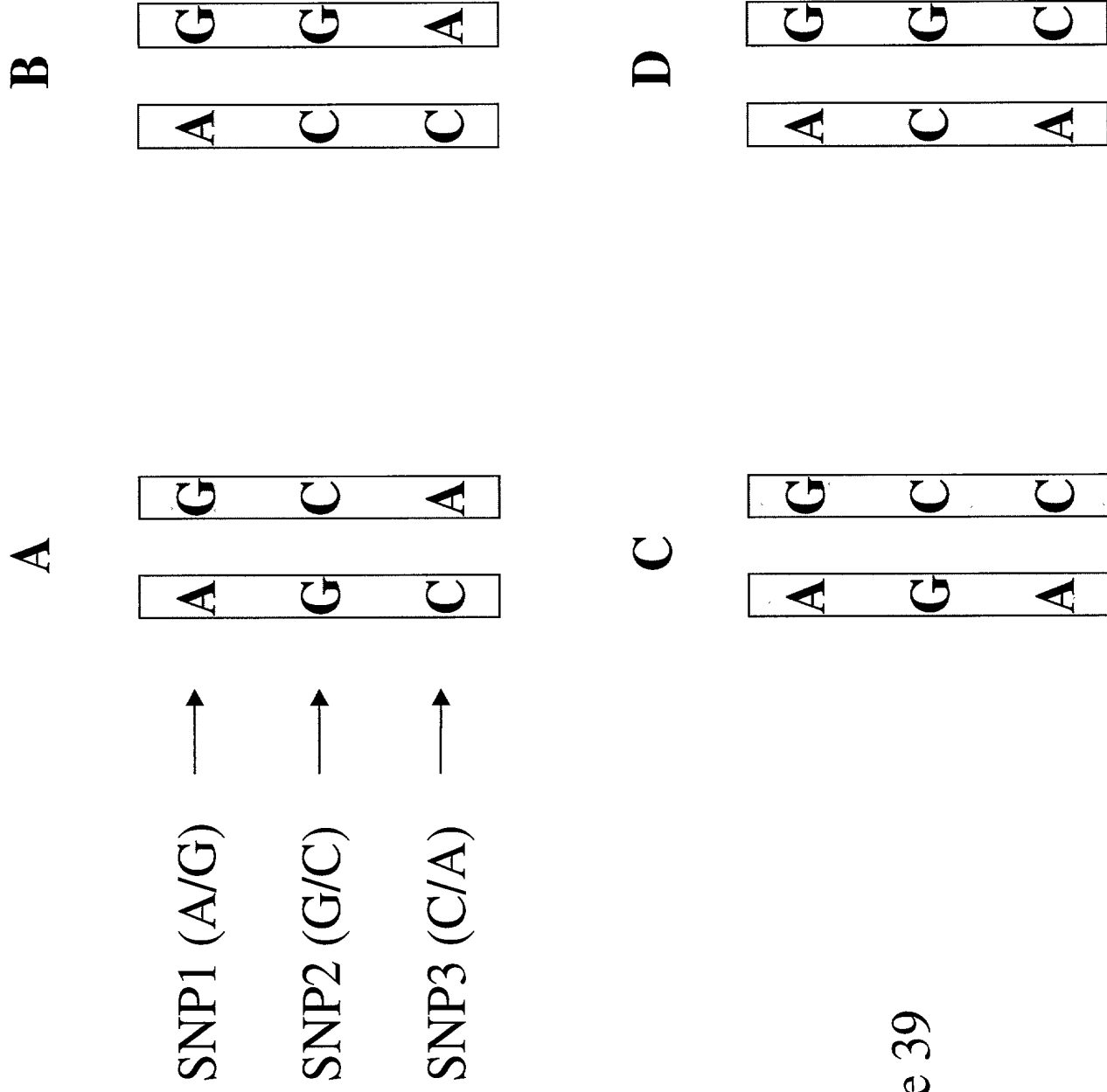


Figure 39

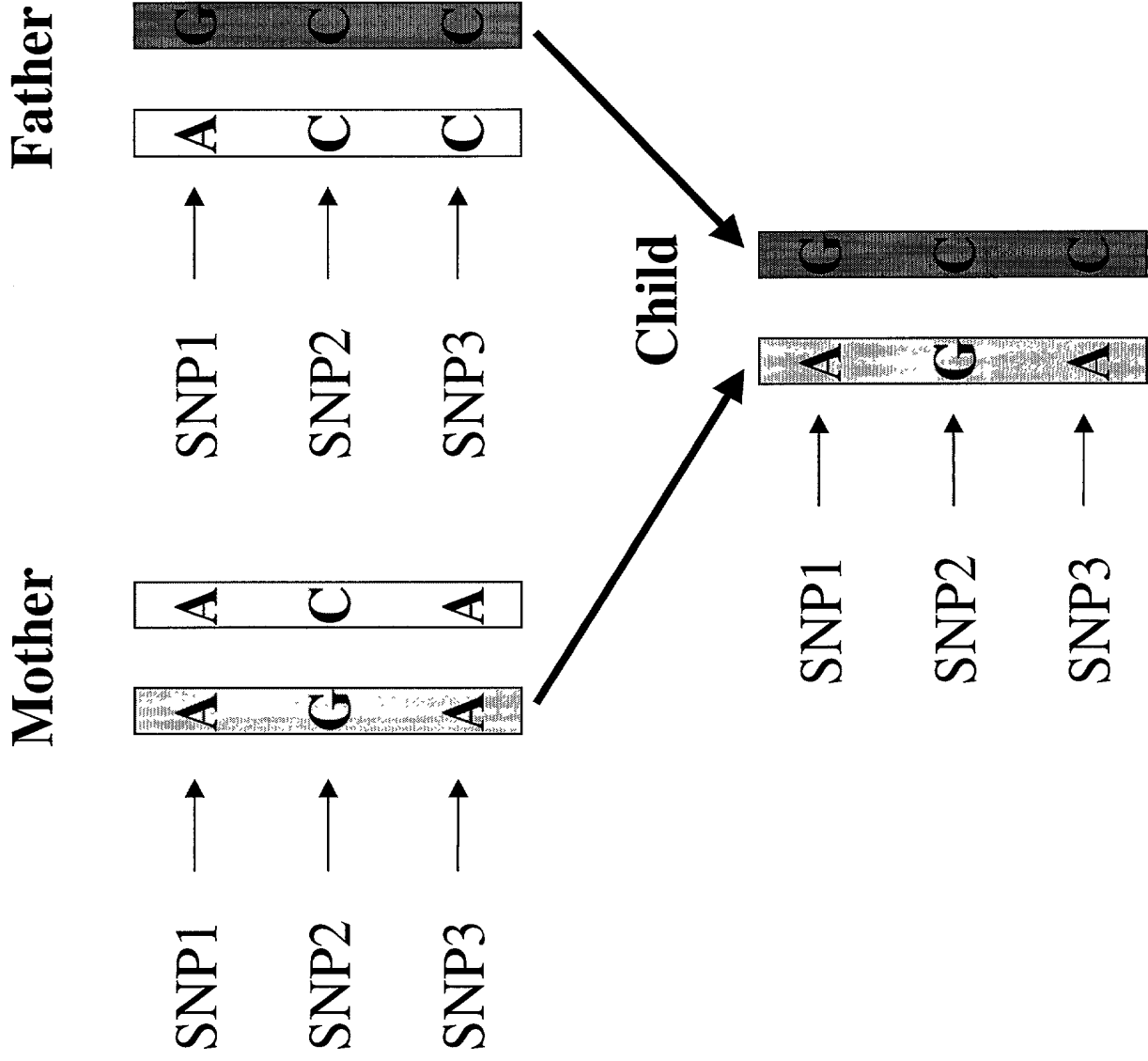


Figure 40

Figure 41A

```

1  CACCCTATCC TACTACTACTA GGAACCTTGCA CAGTCCGCCT CGGGCAGCCC AAAGCTCCTC 60
61  TGCCACCCCT GGCTCCCAA ACCCTCCAAA ACAAAGACC AGAAAAGCAC TCTCCACCCA 120
121 GCAGCCAAAC GCCTCCTTCT TGACGCCAGC CCCACCCCTC TGTCTGCTCG AGCCCAGGAA 180
181 AGGCCTGAAG GAACAGGCCG GGGAAGGAGC CCTCCCTCTC TCCCTTGTC CTCCATCCAC 240
241 CCAGCGCCCG CATCTGGAGA CCCTATGGCC CGGGCTCACT GGGGCTGCTG CCCCTGGCTG 300
1      M A R A H W G C C P W L 12
301 GTCCTCCTCT GTGCTTGTGC CTGGGGCCAC ACAAAGCCAC TGGACCTTGG AGGGCAGGAT 360
13 V L L C A C A W G H T K P L D L G G Q D 32
361 GTGAGAAATT GTTCCACCAA CCCCCCTTAC CTTCCAGTTA CTGTGGTCAA TACCACAATG 420
33 V R N C S T N P P Y L P V T V V N T T M 52
421 TCACTCACAG CCCTCCGCCA GCAGATGCAG ACCCAGAATC TCTCAGCCTA CATCATCCCA 480
53 S L T A L R Q Q M Q T Q N L S A Y I I P 72
481 GGCACAGATG CTCACATGAA CGAGTACATC GGCCAACATG ACGAGAGGCG TGCGTGGATT 540
73 G T D A H M N E Y I G Q H D E R R A W I 92
541 ACAGGCTTTA CAGGGTCTGC AGGAACTGCA GTGGTGAATA TGAAGAAAGC AGCTGTCTGG 600
93 T G F T G S A G T A V V T M K K A A V W 112
601 ACCGACAGTC GCTACTGGAC TCAGGCTGAG CGGCAAATGG ACTGTAATTG GGAGCTCCAT 660
113 T D S R Y W T Q A E R Q M D C N W E L H 132
661 AAGGAAGTTG GCACCACTCC TATTGTCACC TGGCTCCTCA CCGAGATTCC TGCTGGAGGG 720
133 K E V G T T P I V T W L L T E I P A G G 152
721 CGTGTGGGTT TTGACCCCTT CCTCTTGTC ATTGACACCT GGGAGAGTTA TGATCTGGCC 780
153 R V G F D P F L L S I D T W E S Y D L A 172
781 CTCCAAGGCT CTAACAGACA GCTGGTGTCC ATCACAACCA ATCTTGTTGA CCTGGTATGG 840
173 L Q G S N R Q L V S I T T N L V D L V W 192
841 GGATCAGAGA GGCCACCGGT TCCAAATCAA CCCATTTATG CCCTGCAGGA GGCATTACAA 900
193 G S E R P P V P N Q P I Y A L Q E A F T 212
901 GGGAGCACTT GGCAGGAGAA AGTATCTGGC GTCCGAAGCC AGATGCAGAA GCATCAAAAG 960
213 G S T W Q E K V S G V R S Q M Q K H Q K 232

```

Figure 41B

```

961 GTCCCGACTG CCGTCCTTCT GTCGGCGCTT GAGGAGACGG CCTGGCTCTT CAACCTTCGA 1020
233 V P T A V L L S A L E E T A W L F N L R 252

1021 GCCAGTGACA TCCCCTATAA CCCCTTCTTC TATTCCTACA CGCTGCTCAC AGACTCTTCT 1080
253 A S D I P Y N P F F Y S Y T L L T D S S 272

1081 ATTAGGTTGT TTGCCAAACA GAGTCGCTTT AGTCCCGAAA CCTTGAGCTA TCTGAACTCC 1140
273 I R L F A N K S R F S S E T L S Y L N S 292

1141 AGTTGCACAG GCCCCATGTG TGTGCAAATC GAGGATTACA GCCAAGTTCG TGACAGCATC 1200
293 S C T G P M C V Q I E D Y S Q V R D S I 312

1201 CAGGCCTACT CATTGGGAGA TGTGAGGATC TGGATTGGGA CCAGCTATAC CATGTATGGG 1260
313 Q A Y S L G D V R I W I G T S Y T M Y G 332

1261 ATCTATGAAA TGATACCAAG GGAGAAACTC GTGACAGACA CCTACTCCCC AGTGATGATG 1320
333 I Y E M I P R E K L V T D T Y S P V M M 352

1321 ACCAAGGCAG TGAAGAACAG CAAGGAGCAG GCCCTCCTCA AGGCCAGCCA CGTGCGGGAC 1380
353 T K A V K N S K E Q A L L K A S H V R D 372

1381 GCTGTGGCTG TGATCCGGTA CTTGGTCTGG CTGGAGAAGA ACGTGCCCAA AGGCACAGTG 1440
373 A V A V I R Y L V W L E K N V P K G T V 392

1441 GATGAGTTTT CGGGGGCAGA GATCGTGGAC AAGTTCCGAG GAGAAGAACA GTTCTCTCTC 1500
393 D E F S G A E I V D K F R G E E Q F S S 412

1501 GGACCCAGTT TTGAAACCAT CTCTGCTAGT GGTTTGAATG CTGCCCTGGC CCACTACAGC 1560
413 G P S F E T I S A S G L N A A L A H Y S 432

1561 CCGACCAAGG AGCTGAACCG CAAGCTGTCC TCAGATGAGA TGTACCTGCT GGACTCTGGG 1620
433 P T K E L N R K L S S D E M Y L L D S G 452

1621 GGGCAGTACT GGGACGGGAC CACAGACATC ACCAGAACAG TCCACTGGGG CACCCCCTCT 1680
453 G Q Y W D G T T D I T R T V H W G T P S 472

1681 GCCTTTCAGA AGGAGGCATA TACCCGTGTG CTGATAGGAA ATATTGACCT GTCCAGGCTC 1740
473 A F Q K E A Y T R V L I G N I D L S R L 492

1741 ATCTTTCCCG CTGCTACATC AGGGCGAATG GTGGAGGCCT TTGCCCCGAG AGCCTTGTGG 1800
493 I F P A A T S G R M V E A F A R R A L W 512

1801 GATGCTGGTC TCAATTATGG TCATGGGACA GGCCACGGCA TTGGCAACTT CCTGTGTGTG 1860
513 D A G L N Y G H G T G H G I G N F L C V 532

```

Figure 41C

```

1861 CATGAGTGGC CAGTGGGATT CCAGTCCAAC AACATCGCTA TGGCCAAGGG CATGTTCACT 1920
533 H E W P V G F Q S N N I A M A K G M F T 552

1921 TCCATTGAAC CTGGTTACTA TAAGGATGGA GAATTTGGGA TCCGTCTCGA AGATGTGGCT 1980
553 S I E P G Y Y K D G E F G I R L E D V A 572

1981 CTCGTGGTAG AAGCAAAGAC CAAGTACCCA GGGGAGCTAC CTGACCTTGT GGTATCATTT 2040
573 L V V E A K T K Y P G E L P D L V V S F 592

2041 GTGCCCTATG ACCGGAACCT CATCGATGTC AGCCTGCTGT CTCCCAGACA TCTCCAGTAC 2100
593 V P Y D R N L I D V S L L S P E H L Q Y 612

2101 CTGAATCGCT ACTACCAGAC CATCCGGGAG AAGGTGGGTC CAGAGCTGCA GAGGCGCCAG 2160
613 L N R Y Y Q T I R E K V G P E L Q R R Q 632

2161 CTACTAGAGG AGTTCGAGTG GCTTCAACAG CACACAGAGC CCCTGGCCGC CAGGGCCCCA 2220
633 L L E E F E W L Q Q H T E P L A A R A P 652

2221 GACACGCGCT CCTGGGCCTC TGTGTTAGTG GTCTCCACCC TTGCCATCCT TGGCTGGAGT 2280
653 D T A S W A S V L V V S T L A I L G W S 672

2281 GTCTAGAGGC TCCAGACTCT CCTGTTAACC CTCCATCTAG ATGGGGGGCT CCCTTGCTTA 2340
673 V * 673

2341 GCTCCCCTCA CCCTGCACTG AACATACCCC AAGAGCCCCT GCTGGCCCAT TGCCTAGAAA 2400

2401 CCTTTGCATT CATCCTCCTT CTCCAAGACC TATGGAGAAG GTCCCAGGCC CCAGGAAACA 2460

2461 CAGGGCTTCT TGGCCCCAGA TGGCACCTCC CTGCACCCCG GGGTTGTATA CCACACCCTG 2520

2521 GGCCCCTAAT CCCAGGCCCC GAAATAGGAA AGCCAGCTAG TCTCTTCTCT TCTGTGATCT 2580

2581 CAGTAGGCCT AACCTATAAC CTAACACAGA CTGCTACAGC TGCTCCCCTC CCGCCAAACA 2640

2641 AAGCCCCAAG AAAACAATGC CCCTACCACC CAAGGGTGCC ATGGTCCCGG GAAAACCCAA 2700

2701 CCTGTCACCG CGTGTGGGTC GTAACCAGAA CTGTTCCCCC CCACCAGGGC TTAAAAATCG 2760

2761 CCCCCACTTT TTAACCATCG TCCATTAAAC ACCTGGTGGG CATAGCCAGA GCTGTTTCGAA 2820

2821 CCCAGCCAGG GATGAAAAAT CAACCCCCGA CATGGAACCC ATGATTCCTA AACCCGGGGT 2880

2881 AGGTTCCATG CCAAGTAACA GCAGAGGGAG TTAAGCCATA GGAATTTGGC TGTGGAGTAA 2940

```

Figure 41D

2941 GAGGGAATG[·] GGTGAGGCAG[·] TGTGGAATAT[·] GACCCTACCA[·] GAGGTGGAG[·] AACAAACTTG[·] 3000
3001 GGCAGCCGGA[·] ACCCGTCACT[·] ATTTTAGATT[·] CCTGGCATT[·] GAGGAGCCCT[·] TTGAACTTTC[·] 3060
3061 CAAAGTGCAG[·] CCACAGCTAC[·] AATGCTGTTA[·] AATCCTCCCA[·] CATTTCTTGG[·] ATGCCCCCTTC[·] 3120
3121 ACCTTGTGTG[·] GACAGTGTCT[·] GGTTTCCCCA[·] TTTTACAGAC[·] AGGAAAAGT[·] AGCTTCAGAC[·] 3180
3181 AGGGGGTGGG[·] CTTTGCCTAA[·] GGACACACAA[·] ATTTGGTTGG[·] GAGTTGATGG[·] GGCCAGATGA[·] 3240
3241 GCCAGCATT[·] CAGCTGTTTC[·] ACCCTTCAGC[·] AACATGCAGA[·] GTCCCTGAGC[·] CCACCTCCCA[·] 3300
3301 GCCCTCTCCT[·] CATTCTCTGA[·] ACCCACTGTG[·] GTGAGAAGAA[·] TTTGCTCCGG[·] CCAAATTGGC[·] 3360
3361 CGTTAGCCAC[·] CTGGGTCCAC[·] ATCCTGCTAA[·] GACGTTTAAA[·] ACAGCCTAAC[·] AAAGACACTT[·] 3420
3421 GCCTGTGG 3428

Figure 42A

```

1 CTGTGCATGG CATCATCCTG GCCCCCTCTA GAGCTCCAAT CCTCCAACCA GAGCCAGCTC 60
1      M A S S W P P L E L Q S S N Q S Q L 18

61 TTCCCTCAA.ATGCTACGGC CTGTGACAAT GCTCCAGAAG CCTGGGACCT GCTGCACAGA 120
19 F P Q N A T A C D N A P E A W D L L H R 38

121 GTGCTGCCGA CATTATCAT CTCCATCTGT TTCTTCGGCC TCCTAGGGAA CCTTTTTGTC 180
39 V L P T F I I S I C F F G L L G N L F V 58

181 CTGTTGGTCT TCCTCCTGCC CCGGCGGCAA CTGAACGTGG CAGAAATCTA CCTGGCCAAC 240
59 L L V F L L P R R Q L N V A E I Y L A N 78

241 CTGGCAGCCT CTGATCTGGT GTTTGTCTTG GGCTTGCCTT TCTGGGCAGA GAATATCTGG 300
79 L A A S D L V F V L G L P F W A E N I W 98

301 AACCAGTTTA ACTGGCCTTT CGGAGCCCTC CTCTGCCGTG TCATCAACGG GGTCATCAAG 360
99 N Q F N W P F G A L L C R V I N G V I K 118

361 GCCAATTTGT TCATCAGCAT CTTCTGGTGT GTGGCCATCA GCCAGGACCG CTACCGCGTG 420
119 A N L F I S I F L V V A I S Q D R Y R V 138

421 CTGGTGCACC CTATGGCCAG CGGAAGGCAG CAGCGGCGGA GGCAGGCCCG GGTCACCTGC 480
139 L V H P M A S G R Q Q R R R Q A R V T C 158

481 GTGCTCATCT GGGTTGTGGG GGGCCTCTTG AGCATCCCCA CATTCCTGCT GCGATCCATC 540
159 V L I W V V G G L L S I P T F L L R S I 178

541 CAAGCCGTCC CAGATCTGAA CATCACCGCC TGCATCCTGC TCCTCCCCCA TGAGGCCTGG 600
179 Q A V P D L N I T A C I L L L P H E A W 198

601 CACTTTGCAA GGATTGTGGA GTTAAATATT CTGGGTTTCC TCCTACCACT GGCTGCGATC 660
199 H F A R I V E L N I L G F L L P L A A I 218

661 GTCTTCTTCA ACTACCACAT CCTGGCCTCC CTGCGAACGC GGGAGGAGGT CAGCAGGACA 720
219 V F F N Y H I L A S L R T R E E V S R T 238

721 AGAGTGCA.AGG GGCCGAAGGA TAGCAAGACC ACAGCGCTGA TCCTCACGCT CGTGGTTGCC 780
239 R V Q G P K D S K T T A L I L T L V V A 258

781 TTCCTGGTCT GCTGGGCCCC TTACCACTTC TTTGCCTTCC TGGAATTCTT ATTCCAGGTG 840
259 F L V C W A P Y H F F A F L E F L F Q V 278

841 CAAGCAGTCC GAGGCTGCTT TTGGGAGGAC TTCATTGACC TGGGCCTGCA ATTGGCCAAC 900
279 Q A V R G C F W E D F I D L G L Q L A N 298

```

Figure 42B

```
901 TTCTTTGCCT TCACTAACAG CTCCCTGAAT CCAGTAATTT ATGTCTTTGT GGGCCGGCTC 960
299 F F A F T N S S L N P V I Y V F V G R L 318

961 TTCAGGACCA AGGTCTGGGA ACTTTATAAA CAATGCACCC CTAAAAGTCT TGCTCCAATA 1020
319 F R T K V W E L Y K Q C T P K S L A P I 338

1021 TCTTCATCCC ATAGGAAAGA AATCTTCCAA CTTTCTGGC GGAATTAAAA CAGCATTGAA 1080
339 S S S H R K E I F Q L F W R N * 353

1081 CC 1082
```

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 43A

```

1  ATGTTCTCTC  CCTGGAAGAT  ATCAATGTTT  CTGTCTGTTC  GTGAGGCCTC  CGTGCCCAACC  60
1  M  F  S  P    W  K  I    S  M  F    L  S  V  R    E  A  S    V  P  T    20

61  ACGGCCTCTT  TCAGCGCCGA  CATGCTCAAT  GTCACCTTGC  AAGGGCCCAAC  TCTTAACGGG  120
21  T  A  S  F    S  A  D    M  L  N    V  T  L  Q    G  P  T    L  N  G    40

121  ACCTTTGCCC  AGAGCAAATG  CCCCCAAGTG  GAGTGGCTGG  GCTGGCTCAA  CACCATCCAG  180
41  T  F  A  Q    S  K  C    P  Q  V    E  W  L  G    W  L  N    T  I  Q    60

181  CCCCCCTTCC  TCTGGGTGCT  GTTCGTGCTG  GCCACCCTAG  AGAACATCTT  TGTCTCAGC  240
61  P  P  F  L    W  V  L    F  V  L    A  T  L  E    N  I  F    V  L  S    80

241  GTCTTCTGCC  TGCACAAGAG  CAGCTGCACG  GTGGCAGAGA  TCTACCTGGG  GAACCTGGCC  300
81  V  F  C  L    H  K  S    S  C  T    V  A  E  I    Y  L  G    N  L  A    100

301  GCAGCAGACC  TGATCCTGGC  CTGCGGGCTG  CCCTTCTGGG  CCATCACCAT  CTCCAACAAC  360
101  A  A  D  L    I  L  A    C  G  L    P  F  W  A    I  T  I    S  N  N    120

361  TTCGACTGGC  TCTTTGGGGA  GACGCTCTGC  CGCGTGGTGA  ATGCCATTAT  CTCCATGAAC  420
121  F  D  W  L    F  G  E    T  L  C    R  V  V  N    A  I  I    S  M  N    140

421  CTGTACAGCA  GCATCTGTTT  CCTGATGCTG  GTGAGCATCG  ACCGCTACCT  GGCCCTGGTG  480
141  L  Y  S  S    I  C  F    L  M  L    V  S  I  D    R  Y  L    A  L  V    160

481  AAAACCATGT  CCATGGGCCG  GATGCGCGGC  GTGCGCTGGG  CCAAGCTCTA  CAGCTTGGTG  540
161  K  T  M  S    M  G  R    M  R  G    V  R  W  A    K  L  Y    S  L  V    180

541  ATCTGGGGGT  GTACGCTGCT  CCTGAGCTCA  CCCATGCTGG  TGTTCGGAC  CATGAAGGAG  600
181  I  W  G  C    T  L  L    L  S  S    P  M  L  V    F  R  T    M  K  E    200

601  TACAGCGATG  AGGGCCACAA  CGTCACCGCT  TGTGTCATCA  GCTACCCATC  CCTCATCTGG  660
201  Y  S  D  E    G  H  N    V  T  A    C  V  I  S    Y  P  S    L  I  W    220

661  GAAGTGTTC  CCAACATGCT  CCTGAATGTC  GTGGGCTTCC  TGCTGCCCCT  GAGTGTTCATC  720
221  E  V  F  T    N  M  L    L  N  V    V  G  F  L    L  P  L    S  V  I    240

721  ACCTTCTGCA  CGATGCAGAT  CATGCAGGTG  CTGCGGAACA  ACGAGATGCA  GAAGTTCAAG  780
241  T  F  C  T    M  Q  I    M  Q  V    L  R  N  N    E  M  Q    K  F  K    260

781  GAGATCCAGA  CGGAGAGGAG  GGCCACGGTG  CTAGTCCTGG  TTGTGCTGCT  GCTATTTCATC  840
261  E  I  Q  T    E  R  R    A  T  V    L  V  L  V    V  L  L    L  F  I    280

841  ATCTGCTGGC  TGCCCTTCCA  GATCAGCACC  TTCCTGGATA  CGCTGCATCG  CCTCGGCATC  900
281  I  C  W  L    P  F  Q    I  S  T    F  L  D  T    L  H  R    L  G  I    300

```

Figure 43B

```

901 CTCTCCAGCT GCCAGGACGA GCGCATCATC GATGTAATCA CACAGATCGC CTCCTTCATG 960
301 L S S C Q D E R I I D V I T Q I A S F M 320

961 GCCTACAGCA ACAGCTGCCT CAACCCACTG GTGTACGTGA TCGTGGGCAA GCGCTTCCGA 1020
321 A Y S N S C L N P L V Y V I V G K R F R 340

1021 AAGAAGTCTT GGGAGGTGT A CCAGGGAGTG TGCCAGAAAG GGGGCTGCAG GTCAGAACCC 1080
341 K K S W E V Y Q G V C Q K G G C R S E P 360

1081 ATTCAGATGG AGAACTCCAT GGGCACACTG CGGACCTCCA TCTCCGTGGA ACGCCAGATT 1140
361 I Q M E N S M G T L R T S I S V E R Q I 380

1141 CACAACTGC AGGACTGGGC AGGGAGCAGA CAGTGAGCAA ACGCCAGCAG GGCTGCTGTG 1200
381 H K L Q D W A G S R Q * 391

1201 AATTTGTGT A AGGATTGAGG GACAGTTGCT TTTCAGCATG GGCCAGGAA TGCCAAGGAG 1260

1261 ACATCTATGC ACGACCTTGG GAAATGAGTT GATGTCTCCG GTAAACACC GGAGACTAAT 1320

1321 TCCTGCCCTG CCAATTTTG CAGGGAGCAT GGCTGTGAGG ATGGGGTGAA CTCACGCACA 1380

1381 GCCAAGGACT CCAAAATCAC AACAGCATTA CTGTCTTAT TTGCTGCCAC ACCTGAGCCA 1440

1441 GCCTGCTCCT TCCCAGGAGT GGAGGAGGCC TGGGGGAGG GAGAGGAGTG ACTGAGCTTC 1500

1501 CCTCCCGTGT GTTCTCCGTC CCTGCCCCAG CAAGACAACT TAGATCTCCA GGAGAACTGC 1560

1561 CATCCAGCTT TGGTGCAATG GCTGAGTGCA CAAGTGAGTT GTTGCCCTGG GTTTCTTTAA 1620

1621 TCTATTGAGC TAGAACTTTG AAGGACAATT TCTTGCATTA ATAAAGGTTA AGCCCTGAGG 1680

1681 GGTCCCTGAT AACAACTTGG AGACCAGGAT TTTATGGCTC CCCTCACTGA TGGACAAGGA 1740

1741 GGTCTGTGCC AAAGAAGAAT CCAATAAGCA CATATTGAGC ACTTGCTGTA TATGCAGTAT 1800

1801 TGAGCACTGT AGGCAAGACC CAAGAAAGAG AAGGAGCCAT CTCCATCTTG AAGGAACCTA 1860

1861 AAGACTCAAG TGGGAACGAC TGGGCACTGC CACCACCAGA AAGCTGTTCG ACGAGACGGT 1920

1921 CGAGCAGGGT GCTGTGGGTG ATATGGACAG CAGAAGGGGG AGACCAAGGT TCCAGCTCAA 1980

1981 CCAATAACTA TTGCACAACC ACCTGTCCCT GCCTCAGTTC CCTTTTATGT AACATGAAGT 2040

```

Figure 43C

2041 CGTTGTGAGG GTTAAAGGCA GTAACAGGTA TAAAGTACTT AGAAAAGCAA AGGGTGCTAC 2100
 2101 GTACATGTGA GGCATCATTG CGCAGACGTA ACTGGGATAT GTTTACTATA AGGAAAAGAC 2160
 2161 ACTGAGGTCT AGAAATAGCT CCGTGGAGCA GAATCAGTAT TGGGAGCCGG TGGCGGTGTG 2220
 2221 AAGCACCAGT GTCTGGCACA CAGTAGGTGC TCATTGGCTC CCTTCCACCT GTCATTCCCA 2280
 2281 CCACCCTGAG GCCCCAACCG CCACACACAC AGGAGCATTT GGAGAGAAGG CCATGTCTTC 2340
 2341 AAAGTCTGAT TTGTGATGAG GCAGAGGAAG ATATTTCTAA TCGGTCTTGC CCAGAGGATC 2400
 2401 ACAGTGCTGA GACCCCCCAC CACCAGCCGG TACCTGGGAA GGGGAGAGT GCAGGCCTGC 2460
 2461 TCAGGGACTG TTCCTGTCTC AGCAACCAAG GGATTGTTCC TGTCAATCAA TGGTTTATTG 2520
 2521 GAAGGTGGCC CAGTATGAGC CCTAGAAGAG TGTGAAAAGG AATGGCAATG GTGTTACCA 2580
 2581 TCGGCAGTGC CAGGGCAGCA CTCATTCACT TGATAAATGA ATATTTATTA GCTGGTTGGA 2640
 2641 GAGCTAGAAC CTGGAGAGCT AGAACCTGGA GAACTAGAAC CTGGAGGGCT AGAACCTGGA 2700
 2701 GAGGCTAGAA CCAAGAAGGG CTAGAACCTG GAGGGGCTAG AACCTAGAGA AGCTAAAACC 2760
 2761 TGAGCTAGAA GCTGGAGGAC TAGAACCTGG AGGGCTGGAA TCTGAAGGGC TAGAACCTGG 2820
 2821 AGGGCTGGAA TCTGGAGAGC TAGAACCTGG AGGGCTAGAA CCTGGAGGGC TAGAACCTAG 2880
 2881 AAGGGCTAGA ACCTGGAGGG CTGGAATCTG GAGAGCTAGA ACCTGGAGGG CTAGAACCTG 2940
 2941 GAGGGCTAGA ACCTAGAAGG GCTAGAACCT GGAGGGCTAG AACCTGGCAG GTTAGAACCT 3000
 3001 AGAAGGGCTA GAACCTGGAG AGCCAGAACC TGGAGGGCTA GAACCTGGAA GGGCTAGAAC 3060
 3061 CTGTAGAGCT AGAACATGGA GAGCTAGAAC CCGGCAGGCT AGAACCTGGC AAGCTAGAAC 3120
 3121 CTGGAGGGAA TGAACCTGGA GGGCTAGAAC CTGGAGAATG AGAAAAATTT ACATGGCAAA 3180
 3181 GAGCCCATAA ATCTGACCA ATCCAACCTT GAATTTTAAA GCAAAAGCGT GAAAAAAAAG 3240

Figure 43D

3241 ATTCCCTCCT TACCCCAAC CCACTCTTTT TTCCCACCAC CCACTCTCCT CTGCCTCAGT 3300
3301 AAGTATCTGG AGGAAGAAA CAGGTGAAAG AAGAAGTAA AACCATTTAG TATTAGTATT 3360
3361 AGAATGAAGT CAACTGTGC CACACATGGT GAATGAAAA AAAAAAAG AGGCTGTGTT 3420
3421 TTGTCACAC GGCAGTCAT TCAGCACCAG AGCACGTGAT GGTCTGAGAC TCTCTTAGGA 3480
3481 GCAGAGCTCT GCCGCAATGG CCATGTGGGG ATCCACACCT GGTCTGAGGG GCAACTGAGT 3540
3541 CTGCGGGAGA AGAGCGGCC TATGCATGGT GTAGATGCCC TGATAAAGAA CATCTGTCCT 3600
3601 GTGAAAGACT CAATGAGCTG TTATGTTGTA AACAGGAAGC ATTTACATC CAAACGAGAA 3660
3661 AATCATGTAA ACATGTGTCT TTTCTGTAGA GCATAATAAA TGGATGAGGT TTTTGCAAAA 3720
3721 AAAAAAAAA AAA 3733

Figure 44A

1	ATGCATCTTATCGACTACCTGCTCCTCCTGCTGGTTGGACTACTGGCCCTTTCTCATGGC	60
1	M H L I D Y L L L L L V G L L A L S H G	20
61	CAGCTGCACGTTGAGCATGATGGTGAGAGTTGCAGTAACAGCTCCCACCAGCAGATTCTG	120
21	Q L H V E H D G E S C S N S S H Q Q I L	40
121	GAGACAGGTGAGGGCTCCCCAGCCTCAAGATAGCCCCTGCCAATGCTGACTTTGCCTTC	180
41	E T G E G S P S L K I A P A N A D F A F	60
181	CGCTTCTACTACCTGATCGCTTCGGAGACCCCCGGGGAAGAACATCTTTTCTCCCCGCTG	240
61	R F Y Y L I A S E T P G K N I F F S P L	80
241	AGCATCTCGGCGGCCTACGCCATGCTTTCCTGGGGGCCTGCTCACACAGCCGCAGCCAG	300
81	S I S A A Y A M L S L G A C S H S R S Q	100
301	ATCCTTGAGGGCCTGGGCTTCAACCTCACCGAGCTGTCTGAGTCCGATGTCCATAGGGGC	360
101	I L E G L G F N L T E L S E S D V H R G	120
361	TTCCAGCACCTCCTGCACACTCTCAACCTCCCCGCCATGGGCTGGAAACACGCGTGGGC	420
121	F Q H L L H T L N L P G H G L E T R V G	140
421	AGTGCTCTGTTCCTGAGCCACAACCTGAAGTTCCTTGCAAAATTCTGAATGACACCATG	480
141	S A L F L S H N L K F L A K F L N D T M	160
481	GCCGTCTATGAGGCTAAACTCTTCCACACCAACTTCTACGACACTGTGGGCACAATCCAG	540
161	A V Y E A K L F H T N F Y D T V G T I Q	180
541	CTTATCAACGACCACGTCAAGAAGGAAACTCGAGGGAAGATTGTGGATTTGGTCAGTGAG	600
181	L I N D H V K K E T R G K I V D L V S E	200
601	CTCAAGAAGGACGTCTTGATGGTGCTGGTGAATTACATTTACTTCAAAGCCCTGTGGGAG	660
201	L K K D V L M V L V N Y I Y F K A L W E	220
661	AAACCATTCATTTCTCAAGGACCACTCCCAAAGACTTCTATGTTGATGAGAACACAACA	720
221	K P F I S S R T T P K D F Y V D E N T T	240
721	GTCCGGGTGCCCATGATGCTGCAGGACCAGGAGCATCACTGGTATCTTCATGACAGATAC	780
241	V R V P M M L Q D Q E H H W Y L H D R Y	260
781	TTGCCCTGCTCGGTGCTACGGATGGATTACAAAGGAGACGCAACCGTGTTCATTCTC	840
261	L P C S V L R M D Y K G D A T V F F I L	280

Figure 44B

841 CCTAACCAAGGCAAAATGAGGGAGATTGAAGAGGTTCTGACTCCAGAGATGCTAATGAGG 900
281 P N Q G K M R E I E E V L T P E M L M R 300

901 TGGAAACAAC TTGTTGCGGAAGAGGAATTTT TACAAGAAGCTAGAGTTGCATCTTCCCAAG 960
301 W N N L L R K R N F Y K K L E L H L P K 320

961 TTCTCCATTTCTGGCTCCTATGTATTAGATCAGATTTTGCCCAGGCTGGGCTTCACGGAT 1020
321 F S I S G S Y V L D Q I L P R L G F T D 340

1021 CTGTTCTCCAAGTGGGCTGACTTATCCGGCATCACCAAACAGCAAAAAC TGGAGGCATCC 1080
341 L F S K W A D L S G I T K Q Q K L E A S 360

1081 AAAAGTTTCCACAAGGCCACCTTGGACGTGGATGAGGCTGGCACCAGAGGCTGCAGCAGCC 1140
361 K S F H K A T L D V D E A G T E A A A A 380

1141 ACGACGTTTCGCGATCAAAT TCTTCTCTGCCCAGACCAATCGCCACATCCTGCGATTCAAC 1200
381 T T F A I K F F S A Q T N R H I L R F N 400

1201 CGGCCCTTCCTTGTGGTGATCTTTTCCACCAGCACCCAGAGTGTCTCTTTCTGGGCAAG 1260
401 R P F L V V I F S T S T Q S V L F L G K 420

1261 GTCGTCGACCCACGAAACCATAG 1284
421 V V D P T K P * 428

D0053 NP

Figure 45A

1	ATGCATCTTATCGACTACCTGCTCCTCCTGCTGGTTGGACTACTGGCCCTTTCTCATGGC	60
1	M H L I D Y L L L L L V G L L A L S H G	20
61	CAGCTGCACGTTGAGCATGATGGTGAGAGTTGCAGTAACAGCTCCCACCAGCAGATTCTG	120
21	Q L H V E H D G E S C S N S S H Q Q I L	40
121	GAGACAGGTGAGGGCTCCCCAGCCTCAAGATAGCCCCTGCCAATGCTGACTTTGCCTTC	180
41	E T G E G S P S L K I A P A N A D F A F	60
181	CGCTTCTACTACCTGATCGCTTCGGAGACCCCCGGGGAAGAACATCTTTTCTCCCCGCTG	240
61	R F Y Y L I A S E T P G K N I F F S P L	80
241	AGCATCTCGGCGGCCTACGCCATGCTTTCCTGGGGGCCTGCTCACACAGCCGCAGCCAG	300
81	S I S A A Y A M L S L G A C S H S R S Q	100
301	ATCCTTGAGGGCCTGGGCTTCAACCTCACCGAGCTGTCTGAGTCCGATGTCCATAGGGGC	360
101	I L E G L G F N L T E L S E S D V H R G	120
361	TTCCAGCACCTCCTGCACACTCTCAACCTCCCCGGCCATGGGCTGGAAACATGCGTGGGC	420
121	F Q H L L H T L N L P G H G L E T <u>C</u> V G	140
421	AGTGCTCTGTTCCTGAGCCACAACCTGAAGTTCCTTGCAAAATTCCTGAATGACACCATG	480
141	S A L F L S H N L K F L A K F L N D T M	160
481	GCCGTCTATGAGGCTAAACTCTTCCACACCAACTTCTACGACACTGTGGGCACAATCCAG	540
161	A V Y E A K L F H T N F Y D T V G T I Q	180
541	CTTATCAACGACCACGTCAAGAAGGAAACTCGAGGGAAGATTGTGGATTTGGTCAGTGAG	600
181	L I N D H V K K E T R G K I V D L V S E	200
601	CTCAAGAAGGACGTCTTGATGGTGCTGGTGAATTACATTTACTTCAAAGCCCTGTGGGAG	660
201	L K K D V L M V L V N Y I Y F K A L W E	220
661	AAACCATTCAATTCCTCAAGGACCACTCCCAAAGACTTCTATGTTGATGAGAACACAACA	720
221	K P F I S S R T T P K D F Y V D E N T T	240
721	GTCCGGGTGCCCATGATGCTGCAGGACCAGGAGCATCACTGGTATCTTCATGACAGATAC	780
241	V R V P M M L Q D Q E H H W Y L H D R Y	260
781	TTGCCCTGCTCGGTGCTACGGATGGATTACAAAGGAGACGCAACCGTGTTCATTCTC	840
261	L P C S V L R M D Y K G D A T V F F I L	280

[illegible][illegible]

Figure 46A

```

1  CACCCTATCC TACTACTACTA GGAACCTGCA CAGTCCGCCT CGGGCAGCCC AAAGCTCCTC 60
61  TGCCCCACCCT GGCTCCCAAA ACCCTCCAAA ACAAAGACC AGAAAAGCAC TCTCCACCCA 120
121 GCAGCCAAAC GCCTCCTTCT TGACGCCAGC CCCCACCCTC TGTCTGCTCG AGCCCAGGAA 180
181 AGGCCTGAAG GAACAGGCCG GGAAGGAGC CCTCCCTCTC TCCCTTGTCC CTCCATCCAC 240
241 CCAGCGCCGG CATCTGGAGA CCCTATGGCC CGGGCTCACT GGGGCTGCTG CCCCTGGCTG 300
1      M A R A H W G C C P W L 12
301 GTCTCCTCT GTGCTTGTGC CTGGGGCCAC ACAAAGCCAC TGGACCTTGG AGGGCAGGAT 360
13 V L L C A C A W G H T K P L D L G G Q D 32
361 GTGAGAAATT GTTCCACCAA CCCCCCTTAC CTTCAGTTA CTGTGGTCAA TACCACAATG 420
33 V R N C S T N P P Y L P V T V V N T T M 52
421 TCACTCACAG CCCTCCGCCA GCAGATGCAG ACCCAGAATC TCTCAGCCTA CATCATCCCA 480
53 S L T A L R Q Q M Q T Q N L S A Y I I P 72
481 GGACAGATG CTCACATGAA CGAGTACATC GGCCAACATG ACGAGAGGCG TGGTGGATT 540
73 G T D A H M N E Y I G Q H D E R R A W I 92
541 ACAGGCTTTA CAGGGTCTGC AGGAACTGCA GTGGTGACTA TGAAGAAAGC AGCTGTCTGG 600
93 T G F T G S A G T A V V T M K K A A V W 112
601 ACCGACAGTC GCTACTGGAC TCAGGCTGAG CGGCAAATGG ACTGTAATTG GGAGCTCCAT 660
113 T D S R Y W T Q A E R Q M D C N W E L H 132
661 AAGGAAGTTG GCACCACTCC TATTGTCACC TGGCTCCTCA CCGAGATTCC NGCTGGAGGG 720
133 K E V G T T P I V T W L L T E I P A G G 152
721 CGTGTGGGTT TTGACCCCTT CCTCTTGTCC ATTGACACCT GGGAGAGTTA TGATCTGGCC 780
153 R V G F D P F L L S I D T W E S Y D L A 172
781 CTCCAAGGCT CTAACAGACA GCTGGTGTCC ATCACAACCA ATCTTGTGGA CCTGGTATGG 840
173 L Q G S N R Q L V S I T T N L V D L V W 192
841 GGATCAGAGA GGCCACCGGT TCAAATCAA CCCATTTATG CCCTGCAGGA GGCATTACAA 900
193 G S E R P P V P N Q P I Y A L Q E A F T 212
901 GGGAGCACTT GGCAGGAGAA AGTATCTGGC GTCCGAAGCC AGATGCAGAA GCATCAAAAG 960
213 G S T W Q E K V S G V R S Q M Q K H Q K 232

```

Figure 46B

```

961 GTCCCGACTG CCGTCCTTCT GTCGGCGCTT GAGGAGACGG CCTGGCTCTT CAACCTTCGA 1020
233 V P T A V L L S A L E E T A W L F N L R 252

1021 GCCAGTGACA TCCCCTATAA CCCCTTCTTC TATTCCTACA CGCTGCTCAC AGACTCTTCT 1080
253 A S D I P Y N P F F Y S Y T L L T D S S 272

1081 ATTAGGTTGT TTGCAAACAA GAGTCGCTTT AGCTCCGAAA CCTTGAGCTA TCTGAACTCC 1140
273 I R L F A N K S R F S S E T L S Y L N S 292

1141 AGTTGCACAG GCCCATGTG TGTGCAAATC GAGGATTACA GCCAAGTTCG TGACAGCATC 1200
293 S C T G P M C V Q I E D Y S Q V R D S I 312

1201 CAGGCCTACT CATTGGGAGA TGTGAGGATC TGGATGGGA CCAGCTATAC CATGTATGGG 1260
313 Q A Y S L G D V R I W I G T S Y T M Y G 332

1261 ATCTATGAAA TGATACCAAG GGAGAACTC GTGACAGACA CCTACTCCCC AGTGATGATG 1320
333 I Y E M I P R E K L V T D T Y S P V M M 352

1321 ACCAAGGCAG TGAAGAACAG CAAGGAGCAG GCCCTCCTCA AGGCCAGCCA CGTGCGGGAC 1380
353 T K A V K N S K E Q A L L K A S H V R D 372

1381 GCTGTGGCTG TGATCCGTA CTTGGTCTGG CTGGAGAAGA ACGTGCCCAA AGGCACAGTG 1440
373 A V A V I R Y L V W L E K N V P K G T V 392

1441 GATGAGTTT CCGGGGCGAG GATCGTGGAC AAGTCCGAG GAGAAGAACA GTTCTCTCTC 1500
393 D E F S G A E I V D K F R G E E Q F S S 412

1501 GGACCCAGTT TTGAAACCAT CTCTGCTAGT GGTTTGAATG CTGCCCTGGC CCACTACAGC 1560
413 G P S F E T I S A S G L N A A L A H Y S 432

1561 CCGACCAAGG AGCTGAACCG CAAGCTGTCC TCAGATGAGA TGTACCTGCT GGACTCTGGG 1620
433 P T K E L N R K L S S D E M Y L L D S G 452

1621 GGGCAGTACT GGGACGGGAC CACAGACATC ACCAGAACAG TCCACTGGGG CACCCCTCT 1680
453 G Q Y W D G T T D I T R T V H W G T P S 472

1681 GCCTTTCAGA AGGAGGCATA TACCCGTGTG CTGATAGGAA ATATTGACCT GTCCAGGCTC 1740
473 A F Q K E A Y T R V L I G N I D L S R L 492

1741 ATCTTTCCCG CTGCTACATC AGGGCGAATG GTGGAGGCCT TTGCCCAGAG AGCCTTGTGG 1800
493 I F P A A T S G R M V E A F A R R A L W 512

1801 GATGCTGGTC TCAATTATGG TCATGGGACA GGCCACGGCA TTGGCAACTT CCTGTGTGTG 1860
513 D A G L N Y G H G T G H G I G N F L C V 532

```

Figure 46C

1861 CATGAGTGGC CAGTGGGATT CCAGTCCAAC AACATCGCTA TGGCCAAGGG CATGTTCACT 1920
 533 H E W P V G F Q S N N I A M A K G M F T 552

1921 TCCATTGAAC CTGGTTACTA TAAGGATGGA GAATTGGGA TCCGTCTCGA AGATGTGGCT 1980
 553 S I E P G Y Y K D G E F G I R L E D V A 572

1981 CTCGTGGTAG AAGCAAAGAC CAAGTACCCA GGGGAGCTAC CTGACCTTGT GGTATCATT 2040
 573 L V V E A K T K Y P G E L P D L V V S F 592

2041 GTGCCCTATG ACCGGAACCT CATCGATGTC AGCCTGCTGT CTCCNGAGCA TCTCCAGTAC 2100
 593 V P Y D R N L I D V S L L S P E H L Q Y 612

2101 CTGAATCGCT ACTACCAGAC CATCCGGGAG AAGGTGGGTC CAGAGCTGCA GAGGCGCCAG 2160
 613 L N R Y Y Q T I R E K V G P E L Q R R Q 632

2161 CTACTAGAGG AGTTCGAGTG GCTTCAACAG CACACAGAGC CCCTGGCCGC CAGGGCCCCA 2220
 633 L L E E F E W L Q Q H T E P L A A R A P 652

2221 GACACCGCCT CCTGGGCCTC TGTGTAGTG GTCTCCACCC TTGCCATCCT TGGCTGGAGT 2280
 653 D T A S W A S V L V V S T L A I L G W S 672

2281 GTCTAGAGGC TCCAGACTCT CCTGTAAACC CTCCATCTAG ATGGGGGGCT CCCTTGCTTA 2340
 673 V * 673

2341 GCTCCCCTCA CCCTGCACTG AACATACCCC AAGAGCCCCT GCTGGCCCAT TGCCTAGAAA 2400

2401 CCTTTGCATT CATCCTCCTT CTCCAAGACC TATGGAGAAG GTCCCAGGCC CCAGGAAACA 2460

2461 CAGGGCTTCT TGGCCCCAGA TGGCACCTCC CTGCACCCCG GGGTTGTATA CCACACCCTG 2520

2521 GGCCCTAAT CCCAGGCCCC GAAATAGGAA AGCCAGCTAG TCTCTTCTCT TCTGTGATCT 2580

2581 CAGTAGGCCT AACCTATAAC CTAACACAGA CTGCTACAGC TGCTCCCCCT CCGCCAAACA 2640

2641 AAGCCCCAAG AAAACAATGC CCCTACCACC CAAGGGTGCC ATGGTCCCGG GAAAACCCAA 2700

2701 CCTGTCACCG CGTGTGGGC GTAACCAGAA CTGTTCCCCC CCACCAGGGC TTAAAAATCG 2760

2761 CCCCCACTTT TTAACCATCG TCCATTAACC ACCTGGTGGG CATAGCCAGA GCTGTTCGAA 2820

2821 CCCAGCCAGG GATGAAAAAT CAACCCCGA CATGGAACCC ATGATTCTTA AACCCGGGGT 2880

2881 AGGTTCCATG CCAAGTAACA GCAGAGGGAG TTAAGCCATA GGAATTTGGC TGTGGAGTAA 2940

●●●

2941	GAGGGAATGC	GGTGAGGCAG	TGTGGAATAT	GACCCCTACCA	GAGGTTGGAG	AACAAACTTG	3000
3001	GGCAGCCGGA	ACCCGTCACT	ATTTTAGATT	CCTGGCATTC	GAGGAGCCCT	TTGAACTTTC	3060
3061	CAAAGTGCAG	CCACAGCTAC	AATGCTGTTA	AATCCTCCCA	CATTTCTTGG	ATGCCCCTTC	3120
3121	ACCTTGTTGTG	GACAGTGTCT	GGTTTCCCCA	TTTTACAGAC	AGGAAAACCTG	AGCTTCAGAC	3180
3181	AGGGGGTGGG	CTTTGCCTAA	GGACACACAA	ATTTGGTTGG	GAGTTGATGG	GGCCAGATGA	3240
3241	GCCAGCATTTC	CAGCTGTTTC	ACCCTTCAGC	AACATGCAGA	GTCCCTGAGC	CCACCTCCCA	3300
3301	GCCCTCTCCT	CATTCTCTGA	ACCCACTGTG	GTGAGAAGAA	TTTGCTCCGG	CCAAATTGGC	3360
3361	CGTTAGCCAC	CTGGGTCCAC	ATCCTGCTAA	GACGTTTAAA	ACAGCCTAAC	AAAGACACTT	3420
3421	GCCTGTGG	3428					